Fragmentation dynamics of ionised amino acids and neutral clusters of amino acids in the gas phase: a theoretical study

by

Dariusz Grzegorz Piekarski

Universidad Autónoma de Madrid

Madrid, Spain

Directors:
Dr. Sergio Díaz-Tendero Victoria
Prof. Manuel Alcamí Pertejo

Contents

Acknowledgements

Everyone who likes what they're doing may become a great scientist or not. The story of this thesis began more than 5 years ago when Prof. Otilia Mó and Prof. Manuel Yáñez gave me the chance to come to Spain and start this great adventure.

Later on I met not only great scientist but also great human being. I feel privileged to be supervised by him. I am extremely thankful to you Dr. Sergio Díaz-Tendero for supervising this thesis, for your dedicated time, advice and showing your confidence in me.

I would like to assert my honest gratitude to the director of this thesis Prof. Manuel Alcamí and theoretical advisor Prof. Fernando Martín who gave me the opportunity to work with them and learn from everyone in the group: mainly from Dr. Jesús González-Vázquez and Dr. Yang Wang, thank you.

It is also my pleasure to thank Prof. Maria João Ramos, Dr. Enrico Tapavicza and Dr. Patrick Rousseau for sharing your knowledge and experience with me, for three months in your groups in Porto, Long Beach and Caen, respectively.

I would like to thank the whole experimental group in Caen: Arkadiusz Mika, Dr. Rudy Delaunay, Dr. Sylvain Maclot, Dr. Alicja Domaracka, Dr. Patrick Rousseau, Prof. Lamri Adoui and Prof. Bernd Huber for your time, hospitality, full cooperation, delicious home made French food, and shearing nice time eating liquid nitrogen ice creams and drinking glühwein.

I would like to thank Prof. Lorenzo Avaldi, Dr. Paola Bolognesi, Dr. Rita Prosmiti and Dr. Pablo López-Tarifa for numerous scientific discussions.

I would like to thank Prof. Luis Méndez, and Prof. Luis Seijo for sharing time and teaching experience.

Thank you to all my colleagues who have shared lunch-times and countless coffee breaks with me in Spain, Portugal, France and USA.

Dziękuję moim wspaniałym rodzicom, którzy tolerowali moje rzadkie wizyty w domu i których miłość motywuje mnie do działania w każdym aspekcie życia, nie tylko tej pracy doktoranckiej.

Podziękowania należna się również moim kochanym babciom, które zawsze służyły swoimi nieocenionymi radami.

I dla Oli, której miłość uczyniła mnie lepszym człowiekiem.
And to Ola whose love makes me more than a great scientist.

With the support of FPI doctorate grant of the Autonomous University of Madrid and generous allocation of computer time at the Centro de Computatión Científica (CCC).

Abstract

This thesis presents a thorough theoretical study at the molecular level approach of radiation damage of ionised amino acids and derivatives in the gas phase. This thesis only imply the theoretical aspects presented in this manuscript, but in most of the systems the theoretical work has been done in parallel to new experiments carried out at ARIBE, the low-energy ion facility of GANIL (Grand accélérateur national d'ions lourds, Caen, France) and at the GASPHASE beam line of the synchrotron radiation facility ELETTRA (Trieste, Italy). Through all the manuscript theory will be used to interpret the outcomes of experiments in which fragmentation of biomolecules is induced by ion-molecule collisions in the gas phase and analysed by means of mass spectrometry using multicoincidence techniques. In order to have a better understanding of the chemical and physical processes involved, theoretical and experimental results are presented and discussed together.

The present thesis is thus devoted to understand the interaction of ionising radiation with molecules of biological interest. In particular it focuses on the processes taking place just after the irradiation. The radiation damage in biological tissues starts within the first few femtoseconds, during the passage of the impinging particle, at the so-called physical stage. Charge and energy flow leads to ionisation and direct damage. At longer timescales molecular fragmentation occurs (so-called chemical stage of radiation damage), leading to the production of a bunch of possible chemical species, as charged or neutral molecular fragments which also includes the release of high reactive radical species. Thus, the study at the molecular level of the ionisation and fragmentation of the molecular systems is a fundamental step for the understanding and improvement of modern cancer treatments that use beams of protons and other ions, such as C^{6+}. This radiation therapy is called hadrontherapy and represents a more accurate alternative to the irradiation using beams of X rays or accelerated electrons. Moreover, the present studies can also shed light on the formation of molecular species of biological interest in astrophysical environments.

The stability and fragmentation dynamics of several positively charged molecules in the gas phase have been studied: thymidine, glycine, β-alanine, γ-aminobutyric acid, and N-acetylglycine. To face this theoretical challenge, the following computational strategy has been established: the evolution of the ionised and excited molecules in the first hundreds of femtoseconds after the collision, is followed by means of *ab initio* molecular dynamics simulations; analysing the most populated de-excitation channels obtained in these dynamics one gets the molecular fragments produced, from which the relevant parts of the associated potential energy surface is explored using density functional theory. With this approach, the mechanisms, dynamics and energies of the fragmentation channels are obtained together with the intensity and correlation of the fragments observed in the experiments.

The first part of the results includes the evaluation of the energy transferred in an ionising collision. Analysis of the molecular orbitals involved in the charge transfer and the excitation energy distribution is obtained with theoretical simulations. In the second part of the results, the fragmentation

dynamics of doubly charged amino acids is presented. The most important channels correspond to the so-called Coulomb explosion, where the charge is shared into two singly charged fragments, which repeal each other. In competition with this channel, unexpected molecular reorganisation processes are obtained; in particular ultrafast H migration, OH migration and formation of ring molecular structures. These isomerisation processes are typically followed by emission of neutral moieties, leading to stable doubly charged species that are also observed in the experiments. In the case of N-acetylglycine, isomerisation processes (tautomerisation), enabled by the presence of a peptide bond, produce metastable cationic species before undergoing fragmentation. Finally, the last part of the thesis is devoted to understand the structure and stability of molecular clusters formed by β-alanine molecules. This part of the study focuses on the importance of intra- and inter-molecular interactions; in particular it shows that due to the hydrogen bonds strong cooperativity effects are responsible of the extra stability, leading to complex cage-shaped structures, rather than evolving into linear weakly bonded chains. Due to the complexity of these systems, this study is carried out with a different methodological approach: classical molecular dynamics are used to generate large number of conformers and to evaluate their thermal stability; and density functional theory together with wave function analysis methods are used to study different molecular properties: relative stabilities, geometries, bond strengths, dissociation energies, etc.

In summary in this thesis a systematic study on the evolution of ionised amino acids and derivatives based on theoretical simulations combined with mass spectrometry experiments is presented. We report the main factors that govern the stability, fragmentation dynamics and production of stable cationic species after interaction with highly charged ions.

Resumen

Esta tesis presenta un riguroso estudio teórico a nivel molecular del daño producido por radiación en aminoácidos ionizados y sus derivados, en fase gas. En esta tesis únicamente se han desarrollado los aspectos teóricos presentados en el manuscrito, pero en la mayoría de los sistemas estudiados, el trabajo teórico se ha realizado en paralelo a nuevos experimentos llevados a cabo en ARIBE, la instalación de iones de baja energía del GANIL (Grand accélérateur national d'ions lourds , Caen, Francia) y en la línea de haz GASPHASE, de la instalación de radiación de sincrotrón ELETTRA (Trieste, Italia). A través de todo el manuscrito, la teoría será utilizada para interpretar los resultados experimentales en los que la fragmentación de las biomoléculas es inducida por colisiones de las moléculas con iones en fase gas y analizados mediante espectrometría de masas utilizando técnicas de multicoincidencia. Con el fin de tener una mejor comprensión de los procesos químicos y físicos involucrados, los resultados teóricos y experimentales se presentan y discuten al mismo tiempo.

La presente tesis está enfocada a comprender la interacción de la radiación ionizante con moléculas de interés biológico. En particular, se centra en los procesos que tienen lugar justo después de la irradiación. El daño por radiación en tejidos biológicos comienza en los primeros femtosegundos, durante el paso de la partícula que colisiona, en la llamada etapa física. El flujo de carga y energía conduce a la ionización y al daño directo. A intervalos de tiempo más largos se produce la fragmentación molecular (denominada etapa química del daño por radiación), que conduce a la producción de un grupo de especies químicas, como fragmentos moleculares cargados o neutros, y también incluye la liberación de especies radicalarias altamente reactivas. Así, el estudio a nivel molecular de la ionización y fragmentación de los sistemas estudiados, es un paso fundamental para la comprensión y mejora de los tratamientos modernos contra el cáncer que utilizan haces de protones y otros iones, como C^{6+}. Este tipo de radioterapia se denomina hadronterapia y representa una alternativa más precisa a la irradiación utilizando haces de rayos X o electrones acelerados. Además, los estudios que se presentan en este trabajo, también pueden arrojar luz sobre la formación de especies moleculares de interés biológico en ambientes astrofísicos.

Se ha estudiado la estabilidad y la dinámica de fragmentación de varias moléculas cargadas positivamente en la fase gas: timidina, glicina, β-alanina, γ-aminobutírico y *N*-acetilglicina. Para hacer frente a este desafío teórico, se ha establecido la siguiente estrategia computacional: la evolución de las moléculas ionizadas y excitadas en los primeros cientos de femtosegundos después de la colisión, se sigue mediante simulaciones de dinámica molecular *ab initio*; Analizando los canales de desexcitación más poblados obtenidos con estas dinámicas, se obtienen los fragmentos moleculares producidos, a partir de los cuales se exploran las partes más relevantes de la superficie de energía potencial asociada, utilizando la teoría del funcional de la densidad. Con este enfoque, se obtienen los mecanismos, dinámicas y energías de los canales de fragmentación junto con la intensidad y correlación de los fragmentos observados en los experimentos.

La primera parte de los resultados incluye la evaluación de la energía transferida en una colisión ionizante. El análisis de los orbitales moleculares involucrados en la transferencia de carga y la distribución de energía de excitación se han obtenido mediante simulaciones teóricas. En la segunda parte de los resultados se presenta la dinámica de fragmentación de aminoácidos doblemente cargados. Los canales más importantes corresponden a la llamada explosión coulombiana, donde la carga se comparte en dos fragmentos cargados, que se repelen entre sí. En competencia con este canal, se obtienen procesos inesperados de reorganización molecular; En particular migración de H ultrarápida, migración de OH y formación de estructuras moleculares con forma de anillo. Estos procesos de isomerización son típicamente seguidos de la emisión de fragmentos neutros, dando lugar a especies estables y doblemente cargadas, que también son observadas en los experimentos. En el caso de *N*-acetilglicina, los procesos de isomerización (tautomerización), posibilitados por la presencia de un enlace peptídico, producen especies catiónicas metastables antes de sufrir fragmentación. La última parte de la tesis se dedica a comprender la estructura y estabilidad de clusters moleculares formados por moléculas de β-alanina. Esta parte del estudio se centra en la importancia de las interacciones intra- e inter-moleculares; En particular, se muestra que existe una estabilidad adicional debido a los enlaces de hidrógeno, donde se muestran fuertes efectos de cooperatividad, dando lugar a estructuras complejas en forma de jaula, en lugar de evolucionar en cadenas lineales débilmente enlazadas. Debido a la complejidad de estos sistemas, este estudio se lleva a cabo con un enfoque metodológico diferente: Se utilizan simulaciones de dinámica molecular para generar un gran número de confórmeros y evaluar su estabilidad térmica; Y la teoría del funcional de la densidad junto con métodos de análisis de la función de onda se utilizan para estudiar diferentes propiedades moleculares: estabilidades relativas, geometrías, fuerzas de interacción, energías de disociación, etc.

En resumen, en esta tesis se presenta un estudio sistemático sobre la evolución de los aminoácidos ionizados y sus derivados a partir de simulaciones teóricas combinadas con experimentos de espectrometría de masas y se estudian los principales factores que rigen la estabilidad, la dinámica de fragmentación y la producción de especies catiónicas estables, después de la interacción con iones altamente cargados.

Part I

Introduction

The manuscript is divided into four parts. Part I is the *Introduction*, where the context and importance of this work, the basic information about the studied systems, relevant previous works, and a general overview of the computational approach we have followed is presented. *Part II* presents a background information on the theoretical methods that we have applied in this thesis. Part III, entitled *"Results and discussion"* we present in 7 chapters the results obtained for the different systems we have studied. Most of the work has been carried out in close collaboration with experimental groups. Although, this is a theoretical thesis, the experimental results will be also presented for completeness. *Part IV* is a summary of this thesis, including the conclusions we have drawn from our research, suggestions for a future work, and perspectives in the field. Finally, at the end of the thesis we present the bibliography and appendices with supplementary information.

"Nor do I doubt that skilled and scholarly mathematicians will agree with me if, what philosophy requires from the beginning, they will examine and judge, not casually but deeply, what I have gathered together in this book to prove these things."

- Mikołaj Kopernik, precursor of *Scientific Revolution*

Chapter 1

Introduction
Introduction

Theoretical context of this thesis.

This thesis is motivated by the molecular level approach of radiation damage, focusing on the fragmentation processes in a variety of biomolecules, covering different physical phenomena. Combining the experiment and theory opens new perspectives of understanding unexpected reactions. Because of theoretical modelling we are able to study the inaccessible regions of experiments and we can help in the design of new experiments and experimental set-ups. On the other hand, the experimental measurements validate the theoretical calculations, thus the theoretical approaches and modelling can be more effective and adequate. The aim of this thesis is not only to explain the observed fragmentation processes but also to obtain information that can be only reached by the combination of both theoretical modelling and experimental measurements. For instance, coupling of two complement experiments with theoretical simulations allows us to obtain energy-transfer distributions in ionising ion-molecule collisions. This is a quantity that cannot be obtained in the experiments and it is crucial to interpret the experiment and to model the subsequent reactivity of the molecule: i.e. to predict the fragmentation pathways and final products. Information about the final products can be obtained in the experiments, and can be used in turn to validate the theoretical predictions. This interplay between theory and experiment is schematically represented in Fig.1.1 for the particular case of thymidine. As will be presented in Chap.6, dynamic simulations and experimental fragments yields can be combined to obtain energy transfer distributions.

In the last years many scientists have struggled with experimental and financial limitation; due to the fast development of computational science we can use powerful and less expensive computers, and new computational methods, which permit us to do more sophisticated calculations on bigger and bigger systems.[1] Nowadays theoretical chemistry together with computational modelling supported by the supercomputer power have become an integral part of research and development in many fields.[2] In the present thesis the implementation of state-of-the-art quantum

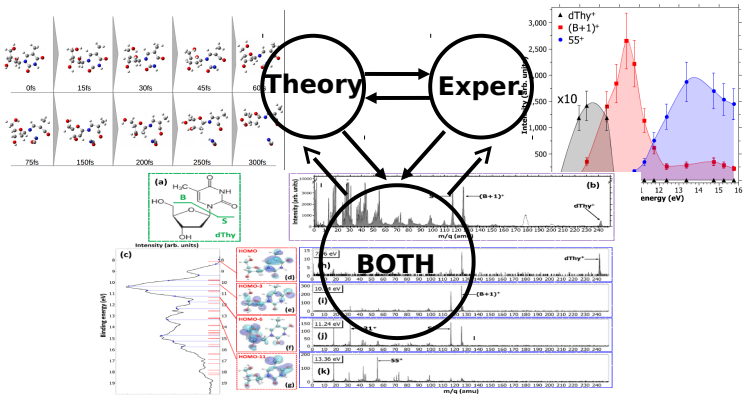


Figure 1.1: The combination of experimental-theoretical scheme applied commonly in this thesis; the example presents the study of the energy distribution in ion-molecule collision (see Chap. 6).

mechanics methods combined with mass spectrometry experiments allows us to study deeply ion-molecule collision phenomena and to understand many aspects of the fragmentation of biomolecules after ionisation in the gas phase.

Before performing any molecular modelling simulation it is essential to answer what our computational limitation is. It depends mainly on the size of the system and the applied methodology. This reasoning makes our study efficient and effective. A reasonable choice of the methodology is a key step to have a detailed knowledge about the fragmentation, reactivity and stability of molecular ions in the gas phase, which is one of the main objectives of this work. Three main strategies, computationally demanding but adequate for our systems, have been applied in this thesis: *ab initio* molecular dynamics (AIMD) simulations, exploration of the Potential Energy Surfaces (PESs), and wave functions analysis methods like Natural Bond Orbitals (NBO) and Quantum Theory of Atoms in Molecules (QTAIM). MD simulations give the kinetic information, fragmentation channels and dissociation mechanisms. Additionally, classical MD simulations performed for neutral clusters provide information of the stability and it is a method itself to generate a wide-range of conformers. An accurate exploration of the PES gives additional energetic and structural information together with minimum energy and fragmentation pathways. Both methodologies facilitate simulating chemical reactions and help to understand in detail the behaviour of the molecular target after the radiation in collisions with highly charged ions. In order to have the whole theoretical picture, tools based on analysis of the wave function have been chosen to describe the interactions in studied systems.

In this thesis we employ these methods to study the behaviour of biomolecules after interaction with multiply charged ions in the gas phase. The biomolecules

chosen are mainly amino acids and its derivatives. We also extended the study of isolated amino acids to understand the structure and stability of clusters of β-alanine. Finally, we aim to extend our investigations and perform simulations for charged clusters and compare the calculations with recent experimental measurements. Ion-molecule/clusters of molecules collisions are highly energetic processes and in many cases the fragmentation can be accompanied by unexpected chemical reactions, which could be properly described by considering only the electronic ground state, therefore by using the density functional theory (DFT). In the case, where DFT fails, we also plan to use linear-response time-dependent DFT and Complete Active Space Self Consistent Field (CASSCF) to complement our study analysing the chemical reactions happening in the electronic excited states such as charge separation due to the fragmentation processes on the non-adiabatic PESs (this is summarised in the perspective chapter).

Thus, in this thesis we present accurate results for a better understanding of mass spectrometry experiments via computer simulation models, as well as the theoretical insight into the nature of the interactions.

Biomolecules under extreme conditions.

The main motivation to study the fragmentation of highly excited/ionised molecules is that they are fundamental processes involved in radiation damage and new cancer treatments, hadrontherapy, based upon the use of charged ions (H^+ or C^{6+}). The full understanding of radiation damage requires knowing which are the physical and chemical steps occurring at molecular level.

In this thesis, we investigate the reactivity of biomolecules induced with an external source (e.g. ionisation and excitation with highly charged ions or photons) at the femtosecond timescale, at which normally these processes occur at living organisms. Ion-biomolecule collisions in the gas phase can be studied under different conditions varying the charge of the projectile (He^{2+}, O^{3+}, O^{6+}, Ar^{9+}, Xe^{25+} etc.) and its collision energy, leading to different physical phenomena like excitation, ionisation, electron capture, charge transfer, Auger decay, etc.

Swift charged ions (H^+ or C^{6+}) interacting with biological tissues deposit most of their energy at a specific depth (the so-called Bragg peak). The release of energy depends on a large extent of the velocity and the charge state of the ions: when the velocity of the ion is much larger than the velocity of the electrons of the target the dominant phenomenon is the ionisation of the target and consequently high energy secondary electrons tracks are produced. When the ion slows down and its velocity becomes of the order of magnitude of the velocity of the electrons of the target molecule, more complex phenomena occur as the capture of electrons by the incident ion and the formation of excited and highly ionised molecules, in these regime the energy transfer from the projectile to the tissue greatly increases leading to the Bragg peak (see Fig.1.2).

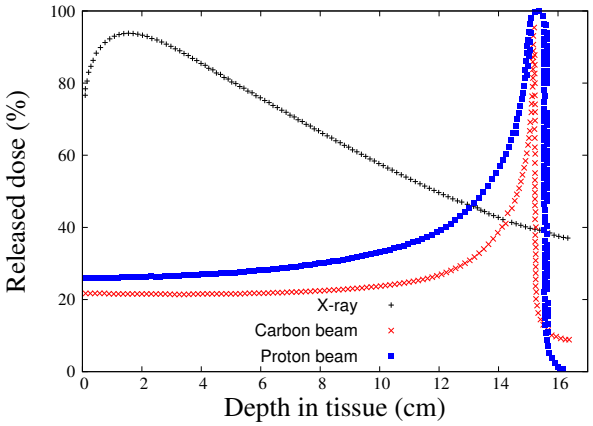


Figure 1.2: The relative dose profile, showing the highest dose release at the end of the path inside of the patient's body, in the region so-called Bragg Peak. The schematic relation of the relative dose in % as a function of the depth in tissue in cm for different radiation sources: X-rays, and beams of different energies (proton and carbon). The maximum penetration of the tumour cell along the track length with minimum side effects is reflected in the narrow C^{6+} beam.

In these complex processes multi-electron capture is a quite probable event, especially when the charge of the projectile is high. Also multiply charge molecules can be produced as a consequence of the Auger cascade produced when a single electron is extracted form inner shells.[3] Therefore one of consequences is that final state of the biomolecule after interaction with the ion beam can be a doubly or triply charged and excited state. From that state the molecule may evolve through many possible chemical reactions: Coulomb explosion, light atom migrations and other unexpected interesting reactions, which take place at the femtosecond timescale. Since they occur in a very short timescale they can be modelled at an *ab initio* level of theory with a reasonable simulation computer time. It is also important to put this timescale in context, these processes occur with respect to the timescale of chemical reactions in living organisms. This is schematically shown in Fig.1.3.

It is also important to indicate that radiation damage is not only produced directly by ion-beam/molecule interactions. High-energy electrons are produced along the beam paths and collides with surounding molecules (indirect effects). Also after ion-biomolecule or ion-water collisions direct ionisation, the subsequent fragmentation chain ends with the formation of small charged ions and radicals that are main factors in the radiation damage.[4–7] Moreover, it has been shown that oxidative reactions of proteins are catalysed by free radicals[8] and also occurred in amino acids, where could damage their structure and in consequence develop diseases like Alzheimer.[9]

To understand these fundamental processes we have focussed our interest in

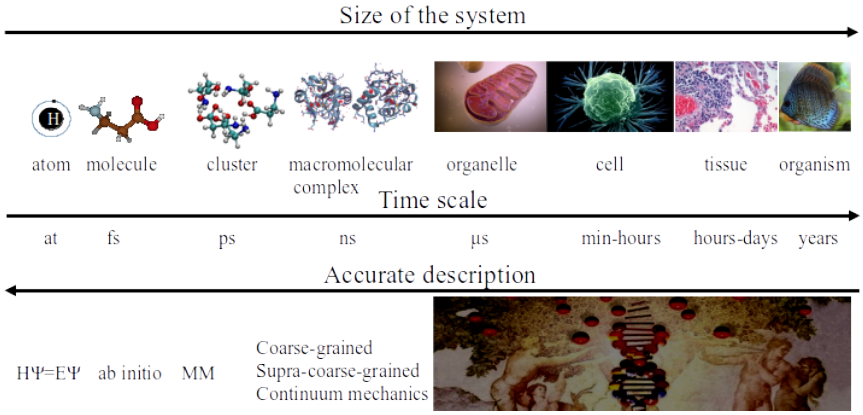


Figure 1.3: Schematic representation of multi-scale processes occurring in matter from the atomic scale to the macroscopic one.

the stability, structure and fragmentation of isolated amino acids, in particular glycine, beta-alanine (β-ala), γ-aminobutiric acid (GABA) (see results shown in Chaps.7-10). We would like to find out what chemical processes can occur in these amino acids, when they are excited and ionised. Amino acids are a specific group of biomolecules. The combinations of only 20 different amino acids bonded by the peptide bond into linear polymer creates proteins (see schematic polypeptide chain with peptide bond in Fig.1.4). Amino acids are the expression of information of life, which is encoded as genetic information in DNA, and a "fingerprint" of the structure of proteins and peptides. Moreover, interestingly amino acids and its precursors also appear in the interstellar space.[10]

To explore in more detail the stability against the fragmentation of the isolated biomolecules we have also chosen the simplest molecule containing the peptide bond -N-acetyl-glycine (see Chap.11).
To determine the distribution of the energy deposited on the ionised molecule as a result of a collision of highly charged ion we have chosen the DNA nucleoside thymidine. The exact value of energy deposited on the molecule for a single collision is unknown due the many reasons. Mainly because (i) the collision with highly charged ions is a very fast process, (ii) occurring at different distances between the projectile and molecule, (iii) with difficult to control distribution of conformers, (iv) as well as it depends on initial temperature equilibration. Thus we can only estimate it by calculation of the distribution as a function of the probability.
The effect of charge localisation in ionic charge states has started to be analysed using the smallest N,N,N-trimethylglycine amino acid - betaine, existing in gas phase exceptionally in zwitterionic form (see Chap.15). In solution and in solid state amino acids exist as zwitterions, where interaction with the environment stabilise them by electrostatic, polarisation, and H-bonding interactions. On the

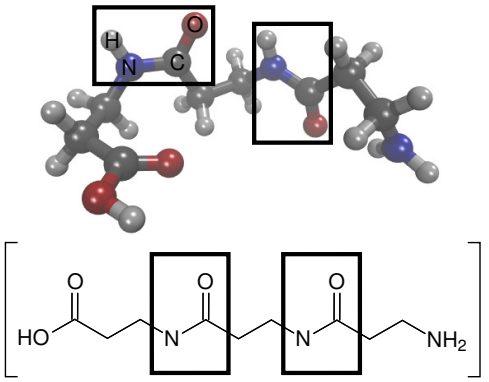


Figure 1.4: Chemical structure of polypeptide with enlighten of two peptide bonds of β-alanine tripeptide. The colours of atoms within the molecules follow the rules: hydrogen (H) is light grey, carbon (C) is dark grey, oxygen (O) is red and nitrogen (N) is blue.

other hand, in the gas phase isolated amino acids do not exist in that form, because intermolecular interactions have no effect.

Although there are fewer isolated β-amino acids than their α counterparts, they are also observed in peptides, in free form and also in the interstellar media.[10] Their geometrical structure i.e. NH_2 terminal position, allows them to form cyclic structures at a relatively low energetic cost e.g. leading to β-lactams which are very reactive biological species.[11;12] Therefore, for the case of neutral β-alanine molecule, we want to understand what happens if, the size of the system increases i.e. when the molecule is embedded into a cluster, to study different types of interactions, the influence of the environment, etc. To do this, we focus on neutral clusters of β-alanine molecules ($(\beta\text{-ala})_n$, where n$\leq$5, see Chap.12). Additional theoretical investigation for charged clusters ($(\beta\text{-ala})_n^{q+}$) was performed allowing comparison with recent experiments based on collisions of molecular clusters with multiply charged ions (see Chap.15).

We have studied the influence of the size (Glycine, alanine, GABA, clusters of β-alanine molecules), charge state (Glycine$^{1+,2+}$), GABA($^{2+,3+}$) and the side chain effect (S-α-alanine vs. β-alanine). The importance of electronic excited states, the presence of a peptide bond and the intermolecular interactions were studied for betaine, N-acetylglycine and clusters of β-alanine, respectively. In Fig.1.5 we present the schematic structure of all molecules discussed in this thesis i.e.: glycine, α- and β-alanine, GABA, N-acetylglycine, betaine and thymidine.

In general, a multiply ionised molecule becomes less stable than its neutral parent due to the extraction of electrons from the bonding regions and the charge-charge repulsion forces acting inside the molecule. The lethal effects to the tissues are also stronger.[13;14] The breaking of the molecule into charged fragments, often referred as Coulomb explosion, leads on many occasions to more stable structures. Therefore only a limited number of molecules are chemically stable as dications in

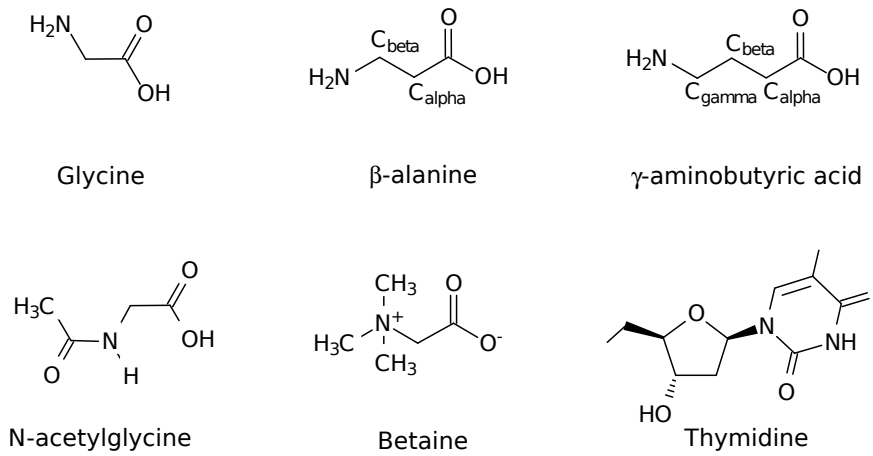


Figure 1.5: Chemical structure of the molecules under studies: glycine, β-alanine, γ-aminobutyric acid, N-acetylglycine, betaine and thymidine.

the gas phase. Moreover, it is important to understand the fragmentation mechanisms and the role of transient species, since multiply charged molecular ions play an important role in the gas phase chemistry.[15–18] Thus, it is very important to get a deep knowledge of structure, stability and fragmentation of amino acids when two or more electrons have been removed.[6;19–21]

Additionally, the study of the conformational behaviour of neutral and ionised amino acids is important for understanding the dynamics of the peptide bond formation. We expect that the exhaustive conformational study as well as the fragmentation analysis presented here, will help to explain the role of the oxidative processes in amino acids and to interpret recent mass spectrometry experiments.

The "cherry" on the cake would be the understanding and explaining of the formation of relevant new species in interstellar space[22] and planet atmospheres.[23] In the gas phase, after the ionisation we can observe ultrafast reactions that can led to species existing in the meteorites. After the meteorite crash in 1969 in Australia a group of scientists performed a careful analysis of its composition. They found that it was rich in organic compounds including amino acids. Besides that, they also found diamino monocarboxylic acids.[24] In CI chondrites (a type of carbonaceous meteorite) β-alanine was confirmed as the most abundant amino acid and its concentration was $\approx$ 10 times higher than for α-alanine.[10] Would they be considered the origin of life on the Earth[25] if they played a key role in space chemistry and were delivered to the early Earth by asteroids, comets, and interplanetary dust particles? The other question is how these compounds can appear in the interstellar space.[26–29] The search of prebiotic molecules, as glycine and β-alanine in the interstellar media, meteorites, and comets has been a hot topic in the last decades centering much experimental[10;30;31] and theoretical efforts.[32–34]

Systems under investigation

Glycine

(NH_2CH_2COOH) is an exceptional amino acid, where the side chain is just one hydrogen atom. The presence in proteins place this amino acid on the 3^{rd} position with 7.2% of abudance. Only alanine and leucine present higher probability to be found in proteins with 8.3% and 9.0% of frequency, respectively.[35] Glycine has been studied by many theoreticians
and experimentalists,[36–47] because it is the simplest amino acid and an important model compound. The first theoretical studies about glycine in gas phase were started in earlies '80.[48;49] They observed the stable enols of radical cation of glycine. From the chemical reactivity point of view it is interesting that the C_α atoms of amino acids can form stable radicals, stabilised by capto-dative effect.[50] For example amyloid-β (A$beta$42) peptide is central to the pathology of Alzheimer disease and includes information about radicals found on C_α of glycine29 and glycine33.[51] The presence of glycine in the interstellar media is still the under debate.[52]

β-alanine

β-alanine is the simplest β-amino acid which in gas phase have been investigated as well by many scientists.[53–56] (β-ala, $NH_2CH_2CH_2COOH$) is a naturally occurring nonessential amino acid. It is a component of co-enzyme A (in pantothenic acid) and carnosine, which has been proven to possess antioxidant properties, it can inhibit the action of free radicals[57–59] and also retard cancer growth in animal models.[60] β-alanine can be used to increase intra-muscular carnosine, which is important in the contraction of muscle tissues and nowadays it is becoming a popular ergogenic aid to sports performance.[61] Increasing the concentration of carnosine in muscles leads to advantages in training adaptations and competition performance arising from its antioxidant properties. Making exercises, hydrogen-ion buildup stimulate muscle's fatigue, and carnosine has been shown to act as a hydrogen-ion buffer[62] in muscle fibers. Carnosine content in muscles is higher for those whose diet is rich in β-alanine.[63] In addition, the presence of this amino acid was found in plants and fruits, as well as in animal brain and insect cuticles.[64] More than that, β-alanine can work as a neurotransmitter in the central nervous system, binding the receptor sides to glycine and γ-aminobutyric acid (GABA)[65] and also plays an important role in the visual system.[66]

γ-aminobutyric acid

($NH_2CH_2CH_2CH_2COOH$, GABA) is the simplest γ-amino acid. Its localisation in living organisms is strictly related to one of the most important type of cells (neurons) i.e. generally speaking is an important molecule in neuroscience.[67–71] Many theoreticians and experimentalists showed the important role played by organic radical cations in various chemical reactions,[58;72–74] where the first step is the ion-

isation of the molecule. Recent studies about radical cations (see e.g. Refs.[75;76]) inspired us to have a closer look at the possible production of multiply charged, highly reactive transient species from GABA, where the evolution of its long chain can also create the cyclic, positively charged, stable and meta-stable radicals.

N-acetylglycine

(aceturic acid) is the simplest N-protected amino acid, where the protecting group is an acetyl group. It represents the smallest system in which one peptide linkage is presented. In the aqueous solution it can create weakly bonded complexes via interaction with transition metals e.g. it has been found that N-acetylglycine coordinates with metal ions such as Cu^{2+}, Cd^{2+}, Zn^{2+}, Ni^{2+} and Co^{2+}.[77] In these systems it has not been observed the deprotonation of the amide hydrogen before the breaking of the metal complex. The presence of the peptide bond makes this molecule worthwhile to investigate its properties in the gas phase after ionisation.

Thymidine

There have been a lot of interest in thymidine as a biochemical tool. It has been long employed as a measure of DNA synthesis, as a cell synchronizing agent in tissue culture and in radioactive form as a cell marker. Its special properties have been evaluated as a "rescue agent" to protect normal tissues from the toxicity induced by methotrexate and fluorodeoxyuridine. Also, it has been reported to posses the anti-cancer activity of cyclophosphamide, and 5-fluorouracil (see Ref.[78] and references within). Our choice to study thymidine seems to be appropriate to reconcile the relevancy of the biological damage, accuracy of the calculations and precision and clarity of the experimental mass spectra.

Clusters of β-alanine

The term of clusters appeared in the late 1970's, when they have drawn a lot of interest. Clusters can be a collection of interacting atoms or molecules. The size of them can vary between a few atoms and aggregates big enough to be close to bulk matter. A common classification of clusters is in small, medium or large size. Its distinction depends on how smoothly the properties of the cluster vary as a function of the cluster size.
In this thesis we present a study of small-size molecular clusters (up to 5 molecules), where the properties as for example the number of intermolecular interactions are changing smoothly but still reflect the small number of component particles. We thus extended the isolated β-alanine molecule study into more complex systems i.e. clusters of β-alanine molecules. Clusters provide an unique environment where the creation of new materials can be triggered, by driving chemical reactions in a new specific way, not possible for isolated molecules. For instances clusters of amino acids can be a perfect environment for the creation of peptide bonds. Clusters may also provide a protective environment against fragmentation when irradiated

with ions.[45] This may give additional information for general chemists and also can be a challenge for theoreticians to provide information of unexpected reactions happened in the cluster environment. Clusters of the same size and exactly the same atomic composition of the residues but different structure could have different chemical reactivity. Size-dependent effects may also be observed in dissociation. Moreover, the calculated energetic information can be directly compared with the experimental measurements.[79;80]
The stability of nanosolvated nucleobasis molecules in water clusters can be very water side-dependent and therefore, very specific isomerisation/fragmentation patterns can occur.[81–91] All this convinced us to choose clusters of β-alanine molecules to performed extensive theoretical study for neutrals and also join the forces with experimentalists and study ionised clusters of β-alanine molecules.

Experimental measurements and interpretation

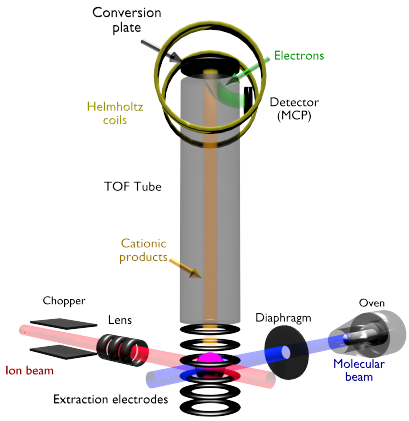


Figure 1.6: Schematic representation of the experimental set-up.*

The ion-molecule collision experiments presented in this thesis were performed at the ARIBE facility of the Grand Accelerateur Nationale of Ion Lourdes (GANIL) in Caen, France. In particular, the COLIMACON experimental setup was used. It is presented schematically in Fig.1.6. The structure of the source of the molecular beam (see details in Fig.1.7) and control of the vapor pressure and the corresponding temperature, allow avoiding thermal decomposition of the molecular targets. The configuration of the electronic aparatus allows to take into account the isotopic contributions. The experimental details can be found in Ref.[92] Only a brief description is given here related to the results shown in this thesis. A beam of projectile ions colliding with a neutral target beam in the gas phase is a very fast

*Thank to Dr. Rudy Delaunay for sharing this experimental figure.

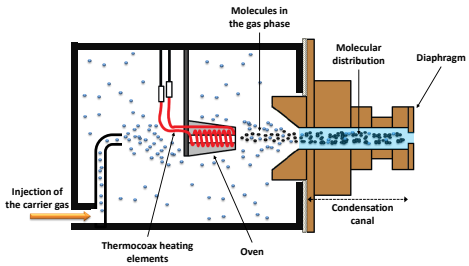


Figure 1.7: Schematic representation of source of the gas phase molecular beam.[†]

process. Therefore, the electrons of the target molecule are rapidly extracted, they can be softly captured but also the molecular target can be ionised and excited simultaneously. Energy from the projectile can be also transferred to the target producing charged and excited species that mainly undergo fragmentation. The probability of the different processes (ionisation, excitation and electron capture) depends on the charge of the projectile and its velocity (see Fig.1.8 taken from Ref.[93]).

Experimental tools were developed to study collisional intermolecular energy transfers in highly energised molecules.[94] From the chemical point of view unimolecular decomposition reaction is conceptually the simplest example. However, the understanding of such decomposition processes has been one of the complex, demanding and controversial debate, even nowadays.[95;96] The analysis of the produced fragments after the ionisation/excitation is a commonly used technique. Positively charged molecular species created in such a way have been analysed with a Wiley-McLaren time-of-flight mass spectrometer. Time-of-flight spectra have been recorded in an event-by-event mode allowing to measure the correlation between the charged fragments proceeding from a single ion-molecule collision by separating the ions as a function of their mass-to-charge (m/z) ratio. The schematic red line in Fig.1.6 represents the low-energy ion beam produced by ion cyclotron resonance and guided through different optical elements. Thus, the ion beam (projectile, P) is focused in the interaction zone called chamber. The blue line represents the molecular neutral beam (target, T), which is prepared in the oven by evaporation of the powder (see Fig.1.7.) The ion beam and the molecular beam overlap in the interaction chamber, thus the negatively charged (Q) electronic cloud of the target interacts with the positively charged (q) projectile and different physical processes occur in competition:

- excitation: $P^{q^+} + T \rightarrow P^{q^+} + T^*$
- electron capture: $P^{q^+} + T \rightarrow P^{q^+ + n_{electrons}} + (T^{-n_{electrons}})^+$

[†]Thank to Dr. Rudy Delaunay for sharing this experimental figure.

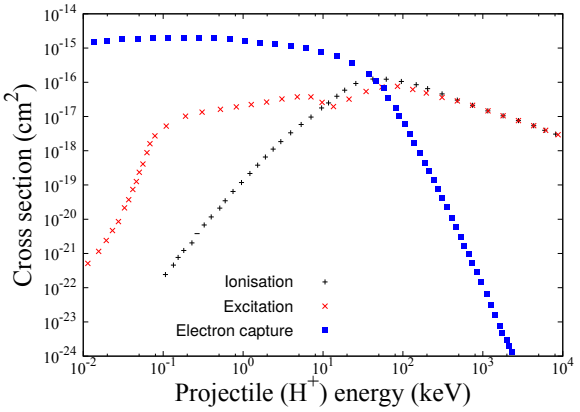


Figure 1.8: Schematic representation of the capture, ionisation and excitation processes as a function of the projectile (P) energy in keV (H^+ + H collision).

- ionisation: $P^{q^+} + T \rightarrow P^{q^+} + (T^{-n_{electrons}})^+ + n_{electrons}$

At lower velocities of the projectile the electron capture dominates over ionisation. At higher velocities, the electron capture becomes the principal process. We are in the experimental region, where the former process is more probable. In summary after ejection of the electron (ionisation, capture, etc.) from the neutral target (isolated molecules, clusters etc.) the ejected electrons or/and the produced charged fragments can be measured. Here we focus on the former case, where we wish to collect additional information about energies, charges and masses of the positively charged species. In practice, we measure, according to the Maxwell's law the mass over charge ratio (m/z). The tries to analyse and distinguish these proceses even for single target atoms as H or He are not a trivial task.[93] These multiple processes for molecular targets (T_M) such as amino acids are given by the equation

$$P^{Q^+} + T_M \rightarrow P^{Q^+ + q_1 n_{electrons}} + (T_M^{-(q_1+q_2)n_{electrons}})^{*+} + q_2 n_{electrons} \qquad (1.1)$$

and are extremely difficult to include in accurate calculations simultaneously, thus the description of some of them is simply neglected. This problem is in a sense of exact solution unsolvable, and approximations need to be applied to theoretical descriptions to correctly predict the physical phenomena occurring in the experiments. In this thesis we are interested in the post collisional processes i.e. the evolution of the charged and excited molecule just after the collision. To this we consider several assumptions. Firstly, we focus only on the target molecule, the ejected electron(s) and the projectile are completely neglected, meaning that the ionisation of the electron(s) happens very fastly in a Franck-Condon type transition. Secondly, we assume that the n electrons are extracted from the highest occupied molecular orbitals. In the case where capture takes place from inner molecular orbitals, we assume that the electronic excitation energy is very rapidly distributed

to nuclear degrees of freedom. The amount of energy transfer from the projectile to the target is included in the term excitation energy and as a whole it is randomly distributed in nuclear velocity.[97] For molecular targets the fragmentation process driven by the collision can be schematically written as:

$$(T_M^{-(q_1+q_2)n_{electrons}})^{*+} \rightarrow [(m_1 + m_2 + \ldots m_n)^{-(q_1+q_2)}]^+ \tag{1.2}$$

, where M is the total mass of the molecule that can fragment into smaller pieces with masses: m_1, m_2 $\ldots m_n$). With the methods detailed in part II of this thesis we study the dynamics of the fragmentation, the different pathways and mechanisms, the structure of the produced fragments and the dissociation energies. Additional experiments have been performed at the GASPHASE beamline of the synchrotron radiation facility ELETTRA in Trieste, Italy. These experiments are also based on crossed-beam set-ups using coincidence time-of-flight mass spectrometry, however here we are able to register the energy of the ejected electrons. The end-station used in the photoionisation experiment consists of a Wiley-McLaren type time-of-flight mass spectrometer coupled to a electron energy analyser in order to perform photoelectron spectroscopy (PS), photoion mass spectrometry (PIMS) and photoelectron-photoion coincidence measurements (PEPICO). The latter measurement stands for multicoincidence detection where the PI (photo-ion) is measured as a function of PE (photo-electron). Based on the electron kinetic energies we can separate different fragmentation channels.[98]

Very accurate description of the processes we study can be achieved combining the theoretical and the experimental methods. This is precisely the main strength of this dissertation.

Part II

Theoretical methods

In 1925, Erwin Schrödinger and Werner Heisenberg independently developed the new quantum theory. The former method involves partial differential equations, whereas the latter method employs matrices. Both methods were shown to be mathematically equivalent and were combined by Paul Dirac into a single formulation. The Schrödinger equation (SE) seems to have a better physical interpretation by means of the classical wave equation.
To obtain the electronic structure and properties of a many-electron system with n electrons and N nuclei it is unavoidable to solve the time-dependent Schrödinger equation (TDSE)

$$\hat{H}(r_i, R_i, t)\Psi(x_i, R_i, t) = i\hbar\frac{\partial}{\partial t}\Psi(x_i, R_i, t), \tag{1.3}$$

where $x_i = (r_i, \eta_i)$ is a space-spin coordinate of the i^{th} electron while $\hat{H}$ depend on the space coordinate of electron r_i, nuclei R_i and time t. If the Hamiltonian does not depend on time, thus the TDSE is simplified to time-independent SE

$$\hat{H}(r_i, R_i)\Psi(x_i, R_i) = E\Psi(x_i, R_i) \tag{1.4}$$

The wave function of the many-body system like a molecule or a cluster is a complex object keeping a large amount of information, which most of the time makes exact or analytical calculations impractical (very often impossible). Therefore, theoreticians need to deal with an appropriate set of approximations specific to the problem, introducing the most effective and accurate methods and computational strategy to solve the SE. In this part all of methods applied in the *"Results and discussion"* are briefly described focusing on two kind of methods: static and dynamic methods. For molecules (in our case 10 atoms for the smallest system - glycine) we have to simplify the SE using the Born-Oppenheimer (BO) approximation,[99] which allows to decouple the electronic and nuclear degrees of freedom.

Chapter 2

Theoretical methods

Electronic structure methods

"Come forth into the light of things, let nature be your teacher."
- William Wordsworth

2.1 Born-Oppenheimer approximation.

The non-relativistic Hamiltonian for a molecule can be written as

$$\hat{H}(r,R) = \hat{T}(r) + \hat{V}(r) + \hat{V}(r,R) + \hat{V}(R) + \hat{T}(R), \tag{2.1}$$

The total Hamiltonian contains the electronic kinetic energy operator

$$\hat{T}(r) = -\frac{\hbar^2}{2m_e}\sum_{i=1}^{n}\nabla_i^2, \tag{2.2}$$

inter electronic Coulomb repulsion

$$\hat{V}(r) = \frac{e^2}{4\pi\epsilon_0}\sum_{i=1}^{n-1}\sum_{j>j}^{n}\frac{1}{|r_i - r_j|}, \tag{2.3}$$

electronic-nuclear Coulomb attraction

$$\hat{V}(r,R) = -\frac{e^2}{4\pi\epsilon_0}\sum_{A=1}^{N}\sum_{i=1}^{n}\frac{Z_A}{|R_A - r_i|}, \tag{2.4}$$

nucleus-nucleus Coulomb repulsion

$$\hat{V}(R) = \frac{e^2}{4\pi\epsilon_0}\sum_{A=1}^{N-1}\sum_{B>A}^{N}\frac{Z_A Z_B}{|R_A - R_B|}, \tag{2.5}$$

and kinetic energy operator for atomic nuclei

$$\hat{T}(R) = -\frac{\hbar^2}{2}\sum_{A=1}^{N}\frac{\nabla_A^2}{M_A}. \tag{2.6}$$

In the above equations the indexes i, j run over all electrons and A, B over all nuclei. m_e and e are the electronic mass and elementary charge of the electron, respectively and ϵ_0 is permittivity of free space; M_A and Z_A are the mass of nucleus and the atomic number of nucleus A, respectively. The Laplacian operators

$$\nabla_i^2 = \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2} \qquad \text{and} \qquad \nabla_A^2 = \frac{\partial^2}{\partial x_A^2} + \frac{\partial^2}{\partial y_A^2} + \frac{\partial^2}{\partial z_A^2} \tag{2.7}$$

act on coordinates of electrons and nuclei, respectively.

Born-Oppenheimer (BO) approximation based on the assumption that the electronic cloud instantaneously adapt to the changes of nuclear configurations. In BO approximation, the $\hat{V}(r,R)$ term can not be neglected, and it prevents us from separation of the total Hamiltonian into nuclear Hamiltonian and electronic Hamiltonian with total wave function Ψ written as a product of electronic ϕ and nuclear χ terms: $\Psi(r,R) = \phi(r)\chi(R)$. However, we can introduce the molecular wave function written as a product of nuclear and electronic wave function, if the latter depends only parametrically on the nuclear configurations R (indicating with semicolon)

$$\Psi(r,R) = \phi(r;R)\chi(R) \tag{2.8}$$

Now, we can fix the nuclear configuration R at some value, and solve for the electronic wave function $\phi(r;R)$, which depends only parametrically on R. Therefore for a fixed nuclear configuration the electronic Hamiltonian can be defined as

$$\hat{H}_{el}(r;R) = \hat{T}(r) + \hat{V}(r) + \hat{V}(r;R), \tag{2.9}$$

We can introduce the so-called effective Hamiltonian, which depends parametrically on nuclear coordinates

$$\hat{H}_{eff}(r;R) = \hat{T}(r) + \hat{V}(r) + \hat{V}(r;R) + \hat{V}(R). \tag{2.10}$$

The main advantage of the BO approximation goes with the assumption that the dependence on R is parametric if the electrons move faster than nuclei then, the statement that the first and the second derivative of the electronic wave function $\phi(r;R)$ with respect to nuclear coordinates are equal to zero must be true:

$$\nabla_A\phi(r;R) \stackrel{!}{=} 0 \qquad and \qquad \nabla_A^2\phi(r;R) \stackrel{!}{=} 0 \tag{2.11}$$

With this assumptions, the general SE in BO approximation for a total wave function from Eq.2.8 is given by

$$\hat{H}(r,R)\Psi(r,R) = \hat{H}\phi(r;R)\chi(R) = E_{tot}\chi(R)\phi(r;R) \tag{2.12}$$

and can be reformulated into two separated Schrödinger equations, electronic:

$$\hat{H}_{eff}(r,R)\phi(r;R) = \left(\hat{T}(r) + \hat{V}(r) + \hat{V}(r;R) + \hat{V}(R)\right)\phi(r;R) = E(R)\phi(r;R) \tag{2.13}$$

and nuclear:

$$\left(\hat{H}_{eff} + \hat{T}(R)\right)\chi(R) = \left(\hat{T}(R) + E(R)\right)\chi(R) \tag{2.14}$$

Solving the electronic SE for a range of R, we obtain the $E(R)$, that act as a potential in Eq.2.14. E(R) is normally referred as potential energy surface (PES) along which the nuclei move. It is customary to replace the nuclear Schrödinger equation with a Newton equation, that is, to move the nuclei classically. In particular for the molecular dynamics (see Chap. 3). The electronic wave function in Eq.2.13 depends on the electron coordinates which depend parametrically only on the nuclear coordinates. This is the so-called "clamped-nuclei" SE. Also it is referred as adiabatic representation, because it neglects the coupling between the electronic states caused by nuclear motion.
Thus the total energy is given by the sum of the electronic energy and the constant nuclear repulsion term:

$$E_{tot} = E + E_{nuc} \tag{2.15}$$

There exist many different implementations to solve Eq.2.13 using some approximations which are different in complexity and in accuracy. We consider the general case in which several electronic excited states are taken into account. If the spectrum of $\hat{H}(r,R)$ is discrete and the eigenfunctions are orthonormal then

$$\int_{-\infty}^{\infty} \phi_i^*(r;R)\phi_j(r;R)dr \equiv \langle\phi_i|\phi_j\rangle = \delta_{ij}. \tag{2.16}$$

In a standard approach the total wave function Ψ may be written in terms of eigenfunctions of $\hat{H}_{eff}$ using an expansion of the form

$$\Psi(r,R,t) = \sum_j \phi_j(r;R)\chi_j(R,t) \tag{2.17}$$

where χ_j is a complete, normalised set of functions which depend parametrically on R. Inserting Eq.2.17 in TDSE

$$\hat{H}(r,t)\Psi(r,R;t) = i\hbar\frac{\partial}{\partial t}\Psi(r,R;t) \tag{2.18}$$

and projecting on the jth state $\phi_j^*(r,R)$, together with integration over only the electronic coordinates gives a set of coupled differential equations

$$[\hat{V}(R) + E_j(R)]\chi_j + \sum_i \hat{\xi}_{ji}\chi_i = i\hbar\frac{\partial}{\partial t}\chi_j \tag{2.19}$$

where $\hat{\xi}_{ji}$ is a coupling operator

$$\hat{\xi}_{ji} \equiv \langle \phi_j | \hat{T}(R) | \phi_i \rangle - \sum_A \frac{\hbar^2}{M_A} \langle \phi_j | \nabla_A | \phi_i \rangle \nabla_A \tag{2.20}$$

The diagonal term ξ_{ii} represents the correction to the adiabatic eigenvalue E_i of the electronic time-independent SE (Eq.2.13). It is proportional to $\frac{1}{M_A}$ to the Born-Oppenheimer potential curve $\hat{V}(R)$ and is very often neglected in the molecular dynamics (MD) and in the surface hoping (SH) methods.[100;101] If also non adiabatic couplings ξ_{ji} are negligible, then the set of differential Eqs.2.19 becomes uncoupled

$$[\hat{V}(R) + E(R)]\chi = i\hbar \frac{\partial}{\partial t} \chi(R). \tag{2.21}$$

In other words, neglecting the second term of Eq.2.19 we obtain the Born-Oppenheimer approximation, in which the nuclear motion is govern by the $\hat{V(R)}$ potential. It means that changing of nuclear coordinates proceed without changes of the quantum state of the electron cloud. Therefore, the wave function from Eq.2.21 in adiabatic representation is given as a single term

$$\Psi(r, R; t) = \phi_i(r, R)\chi_i(R, t) \tag{2.22}$$

One step beyond the Born-Oppenheimer approximation, we retain the off-diagonal coupling operators ξ_{ij} and give foundations to non-adiabatic approximation. This can be govern in MD simulations allowing switches between adiabatic BO PESs and it will be apply in the future (see Chap.15).

Hartree-Fock approximation

The basic method to describe the electronic ground state of the system is the Hartree-Fock[102;103] (HF) method. It is based on the idea that the exact N-body ground state wave function of a system can be expressed approximately by a single Slater determinant Φ_0(in other words, the wave function consists of all antisymmetric products composed of N spin orbitals):

$$\Phi_0 = \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_1(r_1) & \varphi_2(r_1) & \cdots & \varphi_N(r_1) \\ \varphi_1(r_2) & \varphi_2(r_2) & & \varphi_N(r_2) \\ \vdots & \vdots & \ddots & \vdots \\ & & \ddots & \\ \varphi_1(r_N) & \varphi_2(r_N) & \cdots & \varphi_N(r_N) \end{vmatrix} \tag{2.23}$$

which is built up from one-electron wave functions. The pre factor $\frac{1}{\sqrt{N!}}$ ensures normalisation when the one-electron orbitals are orthonormal. The Pauli exclusion

principle is satisfied for a total antisymmetric wave function as the one given by a Slater determinant. If we use Φ_0 to solve the SE imposing that the energy is minimised with respect to variation of orbitals we reach the HF equation given by

$$\hat{F}(r)\Phi_0(r) = E_0\Phi_0(r), \tag{2.24}$$

where

$$\hat{F} = \sum_i^n \hat{f}_i(r), \tag{2.25}$$

and

$$\hat{f}_i(r) = \hat{h}_i(r) + 2\hat{J}_i(r) - \hat{K}_i(r) \tag{2.26}$$

This equation says that $\hat{h}_i$ consists of the kinetic energy of ith electron and its electron-nucleus attraction, Coulomb $\hat{J}_i(r)$ and exchange $\hat{K}_i(r)$ operators characterise the average inter electronic repulsion. They are given by

$$\hat{J}_i(r)\varphi_i(r) = \left[\sum_j^N \int dr' \frac{\varphi_j^*(r')\varphi_j(r')}{|r-r'|}\right]\varphi_i(r) \tag{2.27}$$

and

$$\hat{K}_i(r)\varphi_i(r) = \left[\sum_j^N \int dr' \frac{\varphi_j^*(r')\varphi_i(r')}{|r-r'|}\right]\varphi_j(r). \tag{2.28}$$

The usual way to solve these equations is to expand one-electron function φ_i as linear combination of atomic orbitals-LCAO. Therefore, the one-electron functions φ_i in Eq.2.23 are expanded as:

$$\varphi_i = \sum_\vartheta c_{i\vartheta} f_\vartheta(r; R_i) \tag{2.29}$$

where f_ϑ are the basis functions describing atomic orbitals. The Schrödinger equation for obtaining of the one-electron wave function φ_i is represented as

$$\hat{f}(r_i)\varphi_i(r) = \epsilon_i\varphi_i(r_i), \tag{2.30}$$

Since the eigenvalue is the eigenenergy for the motion of the i^{th} electron, the total wave function can be obtained by solving the eigenequation for each electron iteratively. Such a one-electron wave function and the corresponding eigenenergy are later called the orbital and orbital energy, respectively. The total eigenenergy is the sum of the orbital energies corresponding to different electronic motions. The orbital energy calculated according to Eq.2.24 is given by

$$\epsilon_i = \int \varphi_i^*(r)\hat{F}\varphi_i(r)dr = h_i + \sum_j^n (2J_{ij} - K_{ij}). \tag{2.31}$$

By this the orbital energy the total electron energy is define as

$$E = 2\sum_{i}^{n} \epsilon_i - 2\sum_{i,j}^{n} J_{ij} + \sum_{i,j}^{n} K_{ij} = \sum_{i}^{n} (\epsilon_i + h_i) \tag{2.32}$$

The results of the HF equation is a first approximation to the total energy of the system and a wave function for the ground state expressed as a combination of the molecular orbitals that are also obtained in the calculation.

We have to remember that a single Slater determinant representation excludes Coulomb correlation for electrons with different spins. The electron correlation can be included using either perturbative or variational methods, adding electron correlation through an addition of a small correlation potential to the Hamiltonian of the Hartree-Fock approach or by expressing the total wave function as a linear combination of slater determinants, respectively. The usual implementation of former method is the perturbation theory proposed by Christian Möller and Milton S. Plesset which can be expanded up to n^{th} order (MP2, MP3... MP^n).[104] The latter method is known as the configuration interaction method (CI),[105] and several practical implementation exists such as the quadratic configuration interactions method (QCI)[106] or methods based on Coupled Cluster theories (CC).[107] However, these methods are computationally expensive.[108] The alternative method is the density functional theory (DFT), which includes electron correlation at a low computational cost.

2.2 Density functional theory (DFT)

In the early 1970s, a new electronic structure approach was introduced and named as density functional theory (DFT). The total energy of a system was expressed as a functional of the total electron density. Two theorems postulated by Hohenberg with Kohn and Kohn with Sham proved, the unique relation between electron density and energy, and advance a practical variational DFT approach, respectively.

The Hohenberg-Kohn theorems

The first theorem

In 1964 Hohenberg and Kohn published a paper in Physical Review called *Inhomogeneous Electron Gas*,[109] which is considered the origin of the Density Functional Theory (DFT). They state that:

"*Any observable of a stationary non-degenerated ground state can be calculated, exactly in theory, from the electron density of the ground state. In other words, any observable can be written as a functional of the electron density of the ground state* "

They prof the theorem based on the demonstration that for a given density $\rho(r)$, the external potential $v(r)$ (usually Coulomb potential of the nuclei) is determined

except in an additive constant. This was based on *reductio ad absurdum*, supposing that the form of the theorem is erroneous and leading to conflict.

The exact electronic density $\rho(r)$ is given by:

$$\rho(r) = N \int |\Psi(r, r_2, r_3, \dots r_N)|^2 dr_2 \dots dr_N \tag{2.33}$$

It is associated with the corresponding non-degenerate ground state of N particles (with just one wave function with energy of this state), we can assume that the same electron density can be obtained from two external potentials $\upsilon_1(r)$ and $\upsilon_2(r)$. The potentials differ from each other in more than a constant (when we add a constant to the potential, the wave function and therefore the charge density is unchanged so we have to require that two external potentials are different not only by a constant), what automatically requires generating two different many-electron Hamiltonians $\hat{H}_1 = \hat{T} + \hat{U} + \upsilon_1(r)$ and $\hat{H}_2 = \hat{T} + \hat{U} + \upsilon_2(r)$, with ground state wave functions Ψ_1 and Ψ_2 . $\hat{T}$ in the latter equation is the kinetic energy and $\hat{U}$ is the electron-electron interaction. Thus the two distinct energies are $E_1 = \left\langle \Psi_1 \middle| \hat{H}_1 \middle| \Psi_1 \right\rangle$ and $E_2 = \left\langle \Psi_2 \middle| \hat{H}_2 \middle| \Psi_2 \right\rangle$ respectively.

The variational principle states that the expectation value of the Hamiltonian $(\hat{H})$ of the trial wave function $|\Psi\rangle$ is always equal or greater than the energy of the ground state E_0:

$$\left\langle \Psi \middle| \hat{H} \middle| \Psi \right\rangle = E \geq E_0 = \left\langle \Psi_0 \middle| \hat{H} \middle| \Psi_0 \right\rangle \tag{2.34}$$

The minimum value of $\left\langle \Psi \middle| \hat{H} \middle| \Psi \right\rangle$ is thus an estimation of the variational value of the exact ground state energy. In this way we can calculate the expected value of the energy of Ψ_1 with $\hat{H}_2$:

$$\begin{aligned} E_2 \leq \left\langle \Psi_1 \middle| \hat{H}_2 \middle| \Psi_1 \right\rangle &= \left\langle \Psi_1 \middle| \hat{H}_1 \middle| \Psi_1 \right\rangle + \left\langle \Psi_1 \middle| \hat{H}_2 - \hat{H}_1 \middle| \Psi_1 \right\rangle \\ &= E_1 + \int \rho(r)[\upsilon_2(r) - \upsilon_1(r)]dr \end{aligned} \tag{2.35}$$

Similarly we can also calculate the expected value of the energy of Ψ_2 with $\hat{H}_1$:

$$\begin{aligned} E_1 \leq \left\langle \Psi_2 \middle| \hat{H}_1 \middle| \Psi_2 \right\rangle &= \left\langle \Psi_2 \middle| \hat{H}_2 \middle| \Psi_2 \right\rangle + \left\langle \Psi_2 \middle| \hat{H}_1 - \hat{H}_2 \middle| \Psi_2 \right\rangle \\ &= E_2 + \int \rho(r)[\upsilon_1(r) - \upsilon_2(r)]dr \end{aligned} \tag{2.36}$$

When we summarise Eqs.2.35 and 2.36 we get the inequality:

$$E_1 + E_2 < E_2 + E_1 \tag{2.37}$$

The Eq.2.37 proves that the initial assumption was incorrect: Ψ_1 and Ψ_2 are eigenstates of different Hamiltonians so they are different and because of that the inequality is strict. This demonstrates that it does not exist two different potentials that can be associated to the same ground state electron density. In other words, the ground state electron density can uniquely specify $\upsilon(r)$. Determining the total number of electrons and according the Hohenberg-Kohn theorem (establishing $\upsilon(r)$ by $\rho(r)$), we can say that both the Hamiltonian and the wave function of the ground state are determined by the electron density. In the same way we can also get the expectation value of every ground state variable like, particularly $E = E[\rho]$ or kinetic energy, Coulomb repulsion, etc. We can write the general relation between density and wave function as:

$$\rho(r) \Longrightarrow \upsilon(r) \Longrightarrow \hat{H} \Longrightarrow \Psi \tag{2.38}$$

It is worth to say that the first Hohenberg-Kohn theorem is valid only for non-degenerated ground states, where $\rho(r)$ is N-representable and υ-representable. N-representable means that the ground state electron density has to be a positive function, defined in all space, and its integral has to be equal to the total number of electrons of the system:

$$\rho(r) \geq 0 \qquad \wedge \qquad \int \rho(r)dr = N \tag{2.39}$$

υ-representable refers to the existence of the external potential from which we can derive the electron density. In the case of the trial densities for which we can not find a suitable external potential, inequalities Eq.2.35 and Eq.2.36 can not be achieved. In other cases, the energy is a functional of density:

$$E[\rho] = T[\rho] + U_{Ne}[\rho] + U_{ee}[\rho] \quad (+V_{NN}) \tag{2.40}$$

, where $T[\rho]$ and U_{ee} are universal functionals, independent on the external potential.
Introducing the Hohenberg-Kohn functional $F_{HK}[\rho]$, which usually contains $T[\rho]$ and U_{ee} we can rewrite the Eq.2.40 in terms of the energy, which is a functional of the density for a specific external potential $\upsilon(r)$. Thus, the result from the first Hohenberg-Kohn theorem is:

$$E_\upsilon[\rho] = \int \rho(r)\upsilon(r)dr + F_{HK}[\rho] \tag{2.41}$$

The second theorem

The second theorem supplies the variational principle for $E[\rho]$ and is formulated as follows:

"*The electron density of a non-degenerate ground state can be calculated, exactly in theory, determining the density that minimises the energy of the ground state*".

For a N-representable and v-representable trial density $\tilde{\rho}(r)$ we can write:

$$E_0 \leq E_v[\tilde{\rho}(r)] \tag{2.42}$$

Applying the first Hohenberg-Kohn theorem into the exact and trial densities we will demonstrate the validity of the Eq.2.42. The exact and trial densities define different Hamiltonians and thus, different wave functions.
For the exact density:

$$\rho(r) \rightarrow v_n(r) \rightarrow \hat{H} \rightarrow \Psi \tag{2.43}$$

and for the trial density:

$$\tilde{\rho}(r) \rightarrow \tilde{v}_n(r) \rightarrow \tilde{\hat{H}} \rightarrow \tilde{\Psi} \tag{2.44}$$

Calculating the energy for the trial density, using the Hamiltonian defined by the exact density, we get:

$$E_v[\tilde{\rho}(r)] = \left\langle \tilde{\Psi} \middle| \hat{H} \middle| \tilde{\Psi} \right\rangle \geq \left\langle \Psi \middle| \hat{H} \middle| \Psi \right\rangle = E_0 \tag{2.45}$$

Eq.2.45 is a consequence of the variational principle (Eq.2.34) and is non-equivalent. Moreover, it certifies that for any trial density the resulting energy will be always equal or higher than the exact ground state energy. Looking for a density that minimises the energy, we will find the exact density of the ground state:

$$\frac{\delta E_v[\rho]}{\delta \rho} = 0 \tag{2.46}$$

The density functional theory (DFT)[110] represents an alternative to the conventional high level *ab initio* methods to introduce the effects of electron correlation. The concept of the DFT method is to express the energy of the ground state of a many-electron system E_0, through the electron density ρ, which replace the wave function: $E_0 = E[\rho]$. The main problem appears when we want to find the precise formula that relates energy to the electron density. This formula is not known, so we have to use approximate expressions, which usually provide very good results but, if the functional give wrong results, there is no systematic way to improve it. Despite the difficulties of DFT to deal with open shell multi-configurational electronic states[111] this method is very popular to treat medium and large systems because require less cpu time than high level *ab initio* methods. DFT access the goal of computational chemistry to calculate the energetic properties of chemical

reactions with the accuracy of 1 kcal mol $^{-1}$. DFT methodologies gained the great respect by the chemical community leading to Nobel Prize in Chemistry in 1998 to Walter Kohn.

The Kohn and Sham method

Kohn and Sham proposed a method to calculate the energy from the density by introducing orbitals; strictly it is an expansion of the kinetic energy $T[\Psi]$ for a single determinant function. They proposed a fictitious system for N non-interacting electrons moving under an external potential $\upsilon_s(r)$ and providing a wave function Ψ_s, which has the same density as the real system. One can reproduce the density of the ground state of the real systems by means of studying fictitious systems in which we modified the interaction of the particles. All the quantities related to the fictitious systems will be with index "$_s$". For such a system, excluding electron-electron interactions the orbital approximation provides to the exact HF method, the exact result. For a single-electron term the Hamiltonian is:

$$\hat{H}_s = \sum_{i=1}^{N} \hat{h}_i(r) = \sum_{i=1}^{N} -\frac{1}{2}\nabla_i^2 + \sum_{i=1}^{N} \upsilon_s(r) \tag{2.47}$$

and the exact wave function is:

$$\Psi_s = \frac{1}{\sqrt{N!}} \left|\varphi_1(r_1)\varphi_2(r_2)\varphi_3(r_3)\cdots\varphi_N(r_N)\right| \tag{2.48}$$

Solving the HF equations we can obtain the molecular orbitals:

$$\left[-\frac{1}{2}\nabla_i^2 + \upsilon_s(r)\right]\varphi_i = \epsilon_i\varphi_i \tag{2.49}$$

where, $\langle\varphi_i \mid \varphi_j\rangle = \delta_{ij}$
The exact density (Eq.2.50) and the exact kinetic energy for that system are:

$$\rho(r) = \sum_{i=1}^{N_{occ}} \left|\varphi_i(r)\right|^2 \tag{2.50}$$

$$T_s[\rho] = \sum_{i=1}^{N_{occ}} \left\langle \varphi_i \left| -\frac{1}{2}\nabla_i^2 \right| \varphi_i \right\rangle \tag{2.51}$$

We will get the density of the "real system" by solving the equation:

$$\mu = \upsilon_s(r) + \frac{\delta T_s[\rho]}{\delta\rho(r)} = \frac{\rho E_\upsilon[\rho]}{\delta\rho(r)} \tag{2.52}$$

where

$$E_v[\rho] = \sum_{i=1}^{N_{occ}} \epsilon_i = T_s[\rho] + \int \rho(r) v_s(r) dr \tag{2.53}$$

Introducing the term v_{ee} into Eq.2.53 we will get the real system, in which all N electrons interact:

$$E_v[\rho] = \sum_{i=1}^{N_{occ}} \epsilon_i = T[\rho] + \int \rho(r) v(r) dr + v_{ee} \tag{2.54}$$

Introducing the Coulomb repulsion in terms of the electron density:

$$J[\rho] = \frac{1}{2} \iint \frac{\rho(r_1)\rho(r_2)}{|r_1 - r_2|} dr_1 dr_2 \tag{2.55}$$

we can write the total exchange-correlation energy $E_{XC}[\rho]$ as:

$$E_{XC}[\rho] = (T[\rho] - T_s[\rho]) + (v_{ee}[\rho] - J[\rho]) = T_C[\rho] + W_{XC}[\rho] \tag{2.56}$$

where, $T_C[\rho] = T[\rho] - T_s[\rho]$ and $W_{XC}[\rho] = v_{ee}[\rho] - J[\rho]$
$T_C[\rho]$ is the kinetic correlation energy and $W_{XC}[\rho]$ is the exchange-correlation energy. $E_{XC}[\rho]$ contains all the contributions to the energy that we need to obtain it as a function of the density. We can write it as:

$$E_{XC}[\rho] = \int \rho(r) v_{XC}(r) dr \tag{2.57}$$

Including Eq.2.55 and $E_{XC}[\rho]$ into Eq.2.54 we will get:

$$E_v[\rho] = T[\rho] + \int \rho(r) v(r) dr + \frac{1}{2} \iint \frac{\rho(r_1)\rho(r_2)}{|r_1 - r_2|} dr_1 dr_2 + E_{XC}[\rho] \tag{2.58}$$

Applying in the same way as for Eq.2.52 we have:

$$\mu = \frac{\delta T_s[\rho]}{\delta \rho(r)} + v(r) + \int \frac{\rho(r_2)}{|r_1 - r_2|} dr_2 + \frac{\delta E_{XC}[\rho]}{\delta \rho(r)} = \frac{\rho E_v[\rho]}{\delta \rho(r)} \tag{2.59}$$

If we define the exchange-correlation potential as:

$$v_{XC}(r) = \frac{\delta E_{XC}[\rho]}{\delta \rho(r)} \tag{2.60}$$

and the Coulomb potential as:

$$\phi(r) = v(r) + \int \frac{\rho(r_2)}{|r_1 - r_2|} dr_2 \tag{2.61}$$

we get finally:

$$\mu = \frac{\delta T_s[\rho]}{\delta \rho(r)} + (\phi(r) + v_{XC}(r)) \tag{2.62}$$

or else:

$$\mu = \frac{\delta T_s[\rho]}{\delta \rho(r)} + \upsilon_{eff}(r) \tag{2.63}$$

When we compare Eqs.2.63 with 2.52 we can see that they are the same, when we change $\upsilon_s(r)$ for $\upsilon_{eff}(r)$. This means that the only difference between the real system and the fictitious one lies in the different effective potential $\upsilon_{eff}(r)$, which can be defined as:

$$\upsilon_{eff}(r) = \upsilon_n(r) + \int \frac{\rho(r_2)}{|r_1 - r_2|} dr_2 + \upsilon_{XC}(r) \tag{2.64}$$

Therefore we need to solve the coupled equations:

$$\hat{h}_{KS}\varphi_i = \epsilon_i \varphi_i \tag{2.65}$$

with

$$\hat{h}_{KS} = -\frac{1}{2}\nabla_i^2 + \upsilon_{eff}(r) \tag{2.66}$$

where, $\langle \varphi_i \mid \varphi_j \rangle = \delta_{ij}$. We obtain a set of equations similar like in the HF case. The orbitals in Eqs.2.65 are called Kohn-Sham (KS) orbitals and in terms of them we can calculate the electron density of the ground state:

$$\rho(r) = \sum_{i=1}^{N_{oc}} |\varphi_i(r)|^2 \tag{2.67}$$

Using Eq.2.67 we can similarly to HF apply a procedure for a set of trial molecular orbitals $\{\varphi_i(r)\}$ to obtain the electron density. To solve Eq.2.65 we have to consider a initial set of trial molecular orbitals, to obtain the electron density. From this density we can get $\upsilon_{eff}(r)$ and finally obtain the eigenvalues of Eqs.2.65. This process must be iterative until the convergence criteria will be fulfilled.

Exchange-correlation potential

A system in which the Coulomb repulsion term is included in the potential can not have the same density that the one where it is in not included (the non-interacting one). Adding something that does not interact to make the density the same is done by adding fictitious forces of the nuclei of the molecule. The exchange-correlation potential (υ_{XC}) as introduced in Eq.2.64 is not known and the key issue in DFT is to obtain a good approximation for it. In the next section we briefly summarised the different approaches proposed in the literature. This is still an open question of DFT, and new expression for the XC functional are still under development (see e.g.[112]) and that is because there is not an exact solution so one has to carefully check the selection of the particular functional for the system under study. Therefore some checks of the functionals are recommended i.e. benchmarking the functional for properties of interests. Huge theoretical efforts

were performed to develop a series of families of approximated XC functionals. A good choice of the exchange-correlation potential is a fundamental aspect of DFT. These start with a closed shell system and local density functionals (LDA) based on the known properties of the homogeneous electron gas, moving into the local spin density approximation (LSDA) for open shell systems. Reaching the effects of inhomogeneous densities, both for exchange and correlation, one would use the density gradient for non-local or so-called generalised gradient approximations (GGA). Meta-GGA functionals take into account kinetic energy density (of the Laplacian of density) and hybrid functionals includes the HF exchange.

The most widely used are: local density approximation (LDA), non-local or so-called generalised gradient approximation (GGA), meta-GGA, and hybrid functionals, where B3LYP is commonly known and was mostly used in this thesis.

Local density approximation (LDA)

$E_{XC}[\rho]$ depends particularly on the density. The contribution to the correlation energy is treated independently of the exchange one:

$$E_{XC}^{LDA}[\rho] = E_{X}^{LDA}[\rho] + E_{C}^{LDA}[\rho] \tag{2.68}$$

The analytical expression for the correlation energy based on MP2 perturbation theory has been derived recently[113]

$$E_{C}^{LDA}[r_s] = a\ln(1 + \frac{b}{r_s} + \frac{b}{r_s^2}) \tag{2.69}$$

, where $r_s = (\frac{3}{4\pi\rho})^{\frac{1}{3}}$ is Wigner-Seitz radius equal exactly to one Bohr radius, and a and b are constants evaluated with perturbation theory (see Ref.[113] for details). Then exchange energy per particle with $\alpha = 2/3$ is given as:

$$E_{X}^{LDA}[\rho] = -\frac{9}{4}\alpha\left(\frac{3}{8\pi}\right)^{\frac{1}{3}}\int \rho^{\frac{4}{3}}dr \tag{2.70}$$

and

$$\upsilon_{X}^{LDA}(r) = \frac{\delta(\rho(r)E_X[\rho])}{\delta\rho(r)} = -\frac{3}{2}\alpha\left(\frac{3}{\pi}\right)^{\frac{1}{3}}\rho^{\frac{1}{3}} \tag{2.71}$$

This functional works fine for systems with density which is keep up approximately constant.

Local spin density approximation (LSDA)

α and β electrons have different densities in open shell systems. Because of that for this functional we minimised separately equations for α and β densities. Thus, we can write:

$$E_{XC}^{LSDA}[\rho^{\alpha}] = E_{XC}^{LSDA}[\rho^{\alpha}] + E_{XC}^{LSDA}[\rho^{\alpha}, \rho^{\beta}] \tag{2.72}$$

then:

$$E_{XC}^{LSDA}[\rho^{\alpha}] = -\frac{9}{4}\alpha\left(\frac{3}{8\pi}\right)^{\frac{1}{3}}\int(\rho^{\alpha})^{\frac{4}{3}}\,dr \tag{2.73}$$

and

$$E_{XC}^{LSDA}[\rho^{\alpha},\rho^{\beta}] = \int\rho^{\alpha}(r)E_{C}[\rho^{\alpha},\rho^{\beta}]dr \tag{2.74}$$

As we can see, the energy of α and β are now treated independently, so we have to solve the system of two equations:

$$\delta\left[E_{v}[\rho^{\alpha}] - \mu^{\alpha}\left(\int\rho^{\alpha}(r)dr - N_{\alpha}\right)\right] = 0 \tag{2.75}$$

$$\delta\left[E_{v}[\rho^{\beta}] - \mu^{\beta}\left(\int\rho^{\beta}(r)dr - N_{\beta}\right)\right] = 0 \tag{2.76}$$

Non-local corrections or Generalised Gradient Approximations (GGA)

LDA approximation is based only on the electron density. For GGA we introduce density gradients into the description of exchange and correlation effects. This contributes to an energy functional that depends not only on the density but also on its gradient. The form for a GGA functional becomes:

$$E_{XC}^{GGA}[\rho] = \int f\left(\rho,\nabla\rho\right)dr \tag{2.77}$$

One of the main advantages of GGA, in comparison with LDA, is the significantly better description of the binding energy obtained for molecules. Because of that, DFT-GGA methods gain acceptance in the quantum chemistry community during the early 90's and they are nowadays widely employed.

Meta-GGA

Meta-GGA in comparison with the previous energy functionals, goes even further, including into the description of exchange and correlation the effects of the kinetic energy density $\tau(r)$:

$$\tau(r) = \frac{1}{2}\sum_{i=1}^{N_{occ}} \mid\nabla\varphi_{i}(r)\mid^{2} \tag{2.78}$$

and/or the Laplacian of the density, $\nabla^{2}\rho(r)$:

$$E_{XC}^{meta-GGA}[\rho] = \int f\left(\rho,\nabla\rho,\tau,\nabla^{2}\rho\right)dr \tag{2.79}$$

The presence of the Laplacian requires performing calculations of the second derivatives of the basis functions; larger quadrature grid must therefore be used in calculations.

Hybrid exchange functionals

The existence of the exact connection between the non-interacting density functional system and the fully interacting many body system allows to formulate the exact functional as:

$$E_{XC}[\rho] = \frac{1}{2}\int drdr' \int_{\lambda=0}^{1} d\lambda \frac{\lambda e^2}{|r-r'|}[< \rho(r)\rho(r') >_{\rho,\lambda} - \rho(r)\delta(r-r')] \tag{2.80}$$

where the $< \rho(r)\rho(r') >_{\rho,\lambda}$ is the expectation value of the density-density correlation function and is computed at density $\rho(r)$ for the system described by an effective potential:

$$v_{eff} = v_{en} + \frac{1}{2}\sum_{i\neq j}\frac{\lambda e^2}{|r_i - r_j|} \tag{2.81}$$

Having that and knowing the variation of the density-density correlation function with the coupling constant λ, we can compute the exact energy. When $\lambda = 0$ the non-interacting system corresponds identically to the HF system although the LDA and GGA functionals are constructed to be excellent approximations for the fully interacting system with $\lambda = 1$ (homogeneous electron gas). Then, it is logical to approximate the integral over the coupling constant as a weighted sum of the end points:

$$E_{XC} \approx aE_{FOCK} + bE_{XC}^{GGA} \tag{2.82}$$

, where the a and b coefficients have to be determined by fitting procedures adjusting to reference values (typically experimental measurements).
The most often used energy functional was proposed by Becke, adopted to Eq.2.82 is the Becke three parameters method:[114]

$$E_{XC} = E_{XC}^{LSDA} + a_0\left(E_X^{EXACT} - E_X^{LSDA}\right) + a_x\Delta E_X^{B88} + a_c\Delta E_C^{PW91} \tag{2.83}$$

where, $a_0 = 0.20$, $a_X = 0.72$, $a_C = 0.81$. E_X^{EXACT} is the exact exchange energy, ΔE_X^{B88} is Becke's 1988 gradient correction (to the LSDA) for exchange,[115] and ΔE_C^{PW91} is the 1991 gradient correction for correlation of Perdew and Wang.[110]
Nowadays, hybrid functionals are commonly used in many quantum chemistry calculations. Mostly notable the B3LYP functional is one of the most widely used funcionals in the study of molecules, because parameters were fitted to cover a wide range of molecules with different bonds. Binding energies, geometries, frequencies and other molecular properties obtained with hybrid functionals are generally much reliable than those given by most of the GGA functionals. However, the B3LYP functional sometimes fails as in the case two-center three-electron bonding interactions[116] and it does not properly describe the van der Waals interactions.[117] For neutral clusters of β-alanine molecules (see Chap.F) we used the M06-2X Truhlar high-nonlocality functional with double the amount of non-local exchange (54% of the HF exchange). It has been shown that it properly describe weak interactions in molecular clusters.[118]

Fractional charges in the fragmentation of molecules.

The fragmentation of molecules described by DFT theory with any approximated DFT functional miss the integer nature of electrons. This leads to non-physical fractional occupations of orbitals in a fragment, thus systematic errors in electron density, energy and their derivatives.[119] In our approach, for instance in the trajectories in MD simulations, some bonds are stretched until the molecule is broken, but the DFT calculations is performed over the complete system. In these cases the charges calculated on the fragments will present difficult to interpret non-integer values. Moreover, most of the time we run these calculations for doubly charged systems, where it is important to see in which fragment the charge is located to properly explain the experiments, in which only the fragments with integer charges are measured.

2.3 Basis sets

In the last years a big collection of basis sets was generated in the context of wave function based approaches. The atomic orbitals f_ϑ of the LCAO expansions (see Eq.2.29) are expressed through a set of predefined basis functions $\{\eta_\mu\}$ used to build the approximate wave function. If our aim is to construct a high quality wave function, very large basis sets are necessary. Particularly, basis functions with large values of the angular momentum L (polarisation function) are needed when we want to take into account electron correlation.

Slater Type Orbitals (STO)

Slater type orbitals basis functions are simply exponential $\exp^{-\zeta r}$, which mimic the exact eigenfunctions of the hydrogen atom. Unlike the GTO functions, Slater type orbitals represent the correct cusp behaviour near the nuclei (r$\rightarrow$0) and the the tail regions of wave function (r$\rightarrow \infty$) is properly described (GTO fall off too rapidly). A typically expressed STO is:

$$\eta^{STO} = N r^{n-l} \exp^{-\zeta r} Y_{lm}(\Theta, \phi) \tag{2.84}$$

n is a principal quantum number, ζ corresponds to the orbital exponent and Y_{lm} are the usual spherical harmonics describing the angular part of the function. Despite that STO usually needs three times less functions than GTO to achieve a certain accuracy of describing wave function, there are no analytical methods to compute the two-electron integrals with a STO basis sets.

Gaussian Type Orbitals (GTO)

The set of $\{\eta_\mu\}$ is conventionally chosen as a one consisting of so-called Cartesian Gaussian type orbitals (GTO). Its general form is:

$$\eta^{GTO} = N x^l y^m z^n \exp^{-\alpha r^2} \tag{2.85}$$

N is a normalisation factor ($\langle \eta_\mu \mid \eta_\mu \rangle = 1$, but $\{\eta_\mu\}$ is not orthogonal set; $\langle \eta_\mu \mid \eta_v \rangle \neq 0$ for $\mu \neq v$). α is the orbital exponent and reproduces the compact (large α) or diffuse (small α) behaviour of the function. $L = l + m + n$ is used to classify the GTO as type of functions: s ($L = 0$), p ($L = 1$), d ($L = 2$), etc. the advantage of GTO is that the exponential dependence $\exp^{(-\alpha r^2)}$ allows to solve analytically the four-center-two-electrons integrals. Because of that GTO basis functions are commonly used and more preferable than the Slater type orbitals basis functions.

Contracted Gaussian Function (CGF)

Contracted Gaussian functions are functions (usually between three and six; and not bigger than ten) combined by several primitive Gaussian functions, as the one in Eq.2.85 into a fixed linear combination. In that manner constructed function is called contracted Gaussian function (CGF) and can be written as:

$$\eta_\tau^{GTO} = \sum_a^A d_{a\tau} \eta_a^{GTO} \tag{2.86}$$

We have used in this thesis the density functional theory together with CGF basis sets. This seems the natural choice in Kohn-Sham methods to expand the molecular orbitals.

Minimal basis set

The simplest but also the less accurate expansion of the molecular orbitals uses only one basis functions or one contracted functions for each atomic orbital up to including only the valence orbitals. This type of basis set is called minimal basis set.
The most common representation of this kind of basis set is the STO-3G basis set. It combines three primitive GTO functions into one CGF. For hydrogen only 1s function is included. For carbon there are five functions: two describing the 1s and 2s atomic orbitals and one set of 2p functions (p_x, p_y and p_z), etc. This basis are scarcely used nowadays.

Split valence basis set

Taking into account only the valence space, as the space of changes in the electronic wave function, we can limit the doubled set of functions to the valence orbitals. Therefore, the inner core electrons are treated as in a minimal basis sets. The typical examples are the 3-21G or 6-31G Gaussian basis sets developed by Pople and co-workers. Adding functions with larger angular momentum, p,d.. functions to the hydrogen atoms and d,f.. functions to the chemical elements from the second and third row of the periodic table will augmented the basis set by the polarisation functions. Polarisation functions have more angular nodal planes than the occupied

atomic orbitals. Therefore, the orbitals can distort from their original atomic symmetry and better adapt to the molecular environment. Polarised double zeta or split valence basis sets are necessary to obtain a better description of the long weakly bounded systems (e.g. hydrogen bonds and van der Waals).

6-311++G(d,p)

When we increase the basis set we can obtain triple or quadruple zeta basis sets which are augmented by several sets of polarisation functions including functions (possible higher angular momentum). The cc-pVTZ (for correlation-consistent polarised valence triple zeta) basis sets are typical, modern representatives of this approach in terms of Gaussian functions. Adding extra diffuse functions (sp functions with small exponent α; it means that electrons are kept far away from nuclei) we can enlarge our basis set into 6-311++G(d,p). This basis set use six primitive Gaussians for core atomic orbitals, and the valence orbitals split in three functions each of the them represented by three, one and one Gaussian functions. Diffuse functions "++" are added to the heavy atoms and hydrogen atoms. Polarisation functions: one p set on each hydrogen and d set of each heavy atom are also included. In this manuscript we mostly use the 6-311++G(d,p) basis set in combination with M06-2X or B3LYP functionals. The latter combination was extensively tested.[120]

Basis set superposition error (BSSE)

The concept of fragmentation or dissociation energy (DE) is expressed by interaction (E_{int}) and deformation (E_{def}) energy model in the following manner (see Fig.C.8 and equations within the Chap.F) Although less important than E_{def} and E_{ZPE}, the basis set superposition error (BSSE) has to be included in the proper description of interaction within the clusters of molecules. The size of the basis set is limited by the size of the system. It is known that, for a small basis set, the interaction energy is often too large because a residue uses a part of the basis set of the other residue to improve its own basis set. The error is called the basis set superposition error (BSSE). The problem of BSSE[121;122] is a very well known in literature,[123–125] and appear every time a chemical system split in two (or more components). It is specially important in the case of clusters were the number of residues is high and the interaction energy small, thus the BSSE may seriously falsify the accurate calculations on both hydrogen bond and weakly bound van der Waals clusters.
In 1970, Boys and Bernardi[126] proposed the correction to the BSSE so-called counterpoise (CP) method. The idea of that correction is to calculate the wave functions of the subsystems (residues) and total system (cluster) using the same basis set introducing so-called ghost atoms or raw ghost basis functions. It has been shown that sometimes the CP can overestimate the BSSE[127] due to the effect of the Pauli exclusion principle. For instance for a dimer it is not possible that the fully oc-

cupied orbitals of each monomer are available for occupancy by the electrons of the other monomer. Therefore some scientists claimed that the occupied orbitals should not be available in the monomer calculations and this method is called CP virtual MO's.[128] Many debates and discussions validated both corrections and the conclusion is that both of them seems to provide solid corrections; we use the original full CP correction.

Chapter 3

Theoretical methods
Molecular Dynamics

> "...if we were to name the most powerful assumption of all, which leads one on and on in an attempt to understand life, it is that all things are made of atoms, and that everything that living things do can be understood in terms of the jigglings and wigglings of atoms."
> - Richard Feynman

Simultaneous motion of the atoms and electrons as a molecular unity is commonly known as molecular dynamics (MD). MD simulations became an extremely useful tool to precisely understand a wide range of processes in atoms, molecules, cluster of molecules, solids etc. Depending on the size of the system and the properties that we want to study, different approaches have been developed and implemented. An exact description of the coupled electronic and nuclear degrees of freedom requires solving the time-dependent Schrödinger equation, which is possible only for the simplest systems. [129]

> "At the very heart of any molecular dynamics scheme is the question of how to describe that is in practice how to approximate the interatomic interactions". [130]

In spite of that, certain approximations can be done to deal with the complex molecular systems. As we have seen before, by using of the Born-Oppenheimer approximation the motion of fast and slow degrees of freedom in the MD simulations can be separated. This approach, the so-called adiabatic representation of the molecular system, simply means that the electronic cloud modifies itself to the changes in the nuclear geometry. Accordingly, the potential of the nuclei is calculated by solving the electronic time-independent Schrödinger equation at the discrete nuclear configurations which can be connected into a single potential energy surface (PES), representing single electronic quantum states as a function of the geometries. For most purposes, the lowest electronic energy called ground state is of primary interest. However, there are cases where the Born-Oppenheimer approximation fails. For instance, a charge-transfer or a photo-dissociation processes

involve the crossing of multiple electronic states, which must be included in the dynamics.[131] The non-adiabatic effects can be included into MD simulations by representing the wave function as a set of electronic states coupled throught mechanisms such as non-adiabatic coupling vector, spin-orbit or Renner-Teller.[132;133] The two main approaches found in the literature are called the Ehrenfest method and the surface hopping approach. In the former method the nuclei move on one effective PES, which is an average of all adiabatic states involved, weighted by their populations; it is often called mean-field approximation. The latter method describes the nuclei movement on the pure adiabatic states and switches between them if the population changes.

The MD simulation can be also divided according to the method how the PESs are calculated. The direct and fitted procedures are well known in the literature.[134;135] The direct dynamics approach derives the information as needed on the fly. On the other hand, the fitted methods involve the construction of the entire PES. Both schemes has obviously advantages and disadvantages, but in the fragmentation processes the on the fly strategy is a straightforward choice, because it avoids the imperfect symmetry, degeneracy and fitting errors and only this one will be discussed and applied in this thesis. However, the applicability of this method is very dependent on the electronic structure method used in the dynamics. The calculations of the positions, forces on the nuclei and the gradients at each discrete point, called time step, might exceed thousands for each trajectory, ending up in tens of millions to get a statistical picture of the fragmentation. Moreover, in any of the time step the electronic structure calculation may cause problems with the convergence or lack of energy conservation, thus they need to be systematically inspected during the MD run.

Moreover, further approximations are used to the Born-Oppenheimer, i.e. the MD treats the nuclei as a classical particles and the trajectory is calculated by integrating the equations of motion (Newton's or Lagrange's). Whenever, the electronic part is described by interaction of parametrised mechanical effective potentials or quantum mechanically, the molecular dynamics is called classical or semi classical (*ab initio*), respectively.

This chapter presents as follows: the pure Classical MD based on AMBER code,[136–139] Born-Oppenheimer MD with TURBOMOLE package,[140] further approximation to BOMD, matching into *ab initio* MD (atom-centered density matrix propagation (ADMP)[141;142] and Carr-Parinello MD (CPMD)[143]), nonadiabatic surface hopping linear response time dependent DFT (SH-LR-TDFT)[132] MD with TURBOMOLE and finally nonadiabatic multireference CASSCF MD with surface hopping in the adiabatic representation including arbitrary couplings (SHARC) MD.[133]

3.1 Time integration algorithms

One of the crucial algorithms used in the molecular dynamics simulation is the time integrator algorithm, it integrates the equations of motion of the particles and follows their trajectory. The second most important characteristic of the time integrator is providing the energy conservation on over long periods. Symplectic integrators such as Velocity Verlet algorithm[144;145] implemented in ADMP, excellently satisfy these conditions. However, different integrators are also commonly used. For instance the TURBOMOLE package is using the Leapfrog Verlet[146] integrator algorithm. Time integration algorithms based on finite difference methods are usually a Taylor expansion truncated at some term. Therefore, the time is discretised on a finite grid, thus the time step Δt is the distance between the two sequential points on the grid. Let suppose that positions (R) and their time derivatives at time $t(R)$ are known, thus the integration procedure gives these quantities at next time step ($t+\Delta t$). The time evolution of the system at longer timescales is maintained by applying the iterative scheme. The most popular integration method for MD calculations is the Verlet algorithm presented in the following subsection.

Verlet algorithm

Two variants of the Verlet algorithm are discussed below. The first one is the Velocity Verlet algorithm, previously mentioned and applied in the ADMP dynamics and the second is the Leapfrog Verlet integrator. The general idea of these integrators is two third-order Taylor expansion for the positions $R(t)$, one forward in time:

$$R(t+\Delta t) = R(t) + \dot{R}\Delta t + \frac{1}{2}\ddot{R}(t)\Delta t^2 + \frac{1}{6}\dddot{R}(t)\Delta t^3 + O(\Delta t^4) \tag{3.1}$$

and one backward in time:

$$R(t-\Delta t) = R(t) - \dot{R}\Delta t + \frac{1}{2}\ddot{R}(t)\Delta t^2 - \frac{1}{6}\dddot{R}(t)\Delta t^3 + O(\Delta t^4) \tag{3.2}$$

, where $\dot{R}$ is velocity, $\ddot{R}$ is acceleration, $\dddot{R}$ is the third derivative of R with respect to time. Calculating the sum of two expressions gives the basic form of the Verlet algorithm (also called Störmer's method):

$$R(t+\Delta t) = R(t) - R(t-\Delta t) + \ddot{R}(t)\Delta t^2 + O(\Delta t^4) \tag{3.3}$$

Since we are integrating the Newton's equations, $\ddot{R}(t)$ is just the force divided by the mass, and the force is in turn a function of the positions $R(t)$

$$\ddot{R}(t) = -\frac{1}{M}\nabla V(R(t)) \tag{3.4}$$

A problem with the general form of the Verlet algorithm is that the velocities are not generated directly, even if not needed for the time evolution, they are essential to calculate the kinetic energy, and total energy conservation (kinetic+potential).

Therefore, an additional step has to be performed to compute the velocities from the positions:

$$\dot{R}(t) = \frac{R(t+\Delta t) - R(t-\Delta t)}{2\Delta t} \tag{3.5}$$

Despite the fact that error in Eq.3.3 is of the order of Δt^4, the error associated with the latter equation is of the order of Δt^2. To avoid the latter equation, the Leapfrog and velocity Verlet algorithms were developed. In the latter integration algorithm, the positions, velocities and accelerations at time $t + \Delta t$ are achieved from the same quantities at time t in the following way:

$$R(t+\Delta t) = R(t) + \dot{R}(t)\Delta t + \frac{1}{2}\ddot{R}(t)\Delta t^2 \tag{3.6}$$

$$\dot{R}(t+\frac{\Delta t}{2}) = \dot{R}(t) + \frac{1}{2}\ddot{R}(t)\Delta t \tag{3.7}$$

$$\ddot{R}(t+\Delta t) = -\frac{1}{m}\nabla V(R(t+\Delta t)) \tag{3.8}$$

$$\dot{R}(t+\Delta t) = \dot{R}(t+\frac{\Delta t}{2}) + \frac{1}{2}\ddot{R}(t+\Delta t)\Delta t \tag{3.9}$$

Molecular dynamics simulations are solved numerically; the classical equations of motions are integrated step-by-step; they can be defined for a molecular systems as:

$$M_A\ddot{R}_A = f_A \qquad\qquad f_A = -\frac{\partial}{\partial R_A}V(R^N) \tag{3.10}$$

where we consider a system of N particles moving under the influence of a potential function V; f_A is the forceacting on the A^{th} atom, derived from a potential energy $V(R^N)$, and $R^N = (R_1, R_2, \dots R^N)$ represents the complete set of $3N$ atomic coordinates.
According to the Hamiltonian:

$$\hat{H}(R^N, p^N) = \sum_{i=1}^{N}\frac{p_A^2}{2M_A} + V(R^N) \tag{3.11}$$

we can write Eqs.3.10 in the form:

$$\dot{R}_A = \frac{\partial \hat{H}}{\partial p_A} = \frac{p_A}{M_A} \qquad\qquad \dot{p}_A = -\frac{\partial \hat{H}}{\partial R_A} = -\frac{\partial V}{\partial R_A} = f_A(R^N) \tag{3.12}$$

where R is the position of the nuclei with momenta $p = MV(R^N)$.
The molecular integration algorithm for the nuclei movement can be solved considering Newton's mechanics. Using $p_A = M_A\dot{R}_A$ we can write the Newton's second

law:

$$M_A \ddot{R}_A = f_A(R^N), \tag{3.13}$$

Movement of the nuclei can be also described using the Lagrangian formalism:

$$\mathcal{L}(R^N, \dot{R}^N) = \sum_{A=1}^{N} \frac{1}{2} M_A \dot{R}_A^2 - V(R^N) \tag{3.14}$$

and solving the Euler-Lagrange equations of motion:

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{R}_A} = \frac{\partial \mathcal{L}}{\partial R_A} \tag{3.15}$$

Both formalisms are equivalent, where the second one is more commonly used in *ab initio* molecular dynamics. The Lagrangian technique is used in the CPMD and ADMP MD methods employed in this thesis.

3.2 Classical molecular dynamics

Molecular Mechanics (MM) methods are based on the laws of classical mechanics for prediction of the structure and properties of molecular systems. MM is characterised by a force field,which describes only interactions between nuclei, and electronic part is neglected. Because of that, the electronic effects are included in the energy function. Despite of these limitations there is a wide field of use with this level of theory due to its computational efficiency, which allows modelling of extremely big systems. MM can not be applied to the problems strictly depended on the electronic structure, like fragmentation reactions. Other disadvantages at this level of theory are the force field limitations to describe the typical classes of molecule, mainly neutral one. Classical molecular dynamics use predefined potentials, either based on empirical data or on independent electronic structure calculations. It is a powerful tool to investigate interaction of big many-body systems, like clusters of β-alanine (see Chap.12). Classical molecular dynamics simulations were performed using the AMBER[136–139] package. This is the general name of a bunch of programs created mainly to perform molecular dynamics calculations on biomolecules. AMBER posses a substructure database, force field parameters files and many useful programs as antechamber, leap, ptraj, etc. (the three listed were used in this thesis). These sub programs work together with reasonably good efficiency and became a powerful framework for running calculations of the dynamics at the pico or nano-timescale and even at the microsecond timescale, on big systems.

General Amber Force Field (GAFF)

Force fields are the foundations of molecular mechanics; it is therefore important to choose the one that correctly describe our system without any missing force field parameters. The Antechamber program, introduced into AMBER was written by Junmei Wang, and it was provided with the general AMBER force field (GAFF).[147] GAFF was created mainly to investigate rational drug design but works fine with almost all the organic molecules build of C, N, O, H, S, P, F, Cl, Br, so it should be the right choice for clusters of β-alanine as well. The energy function used in GAFF is a simple function of the form:

$$E_{pair} = \sum_{bonds} K_r(r - r_{eq})^2 + \sum_{angles} K_\theta(\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{V_n}{2}[1 + \cos(n\phi - \gamma)] + \sum_{i<j}[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\varepsilon R_{ij}}] \tag{3.16}$$

Where r_{eq} and q_{eq} are equilibration structural parameters, K_r, K_Θ, V_n are force constants, n is the multiplicity and γ is the phase angle for torsional angle parameters. The A , B , and q parameters characterise the non-bonded potentials. The charge method used in GAFF is HF/6-31G(d,p) RESP charge.[148] Detailed information about algorithms used to classify atoms, bond types and charges assignation to estimate the force field parameters can be found in the paper of Wang et al.[147]

3.3 *Ab initio* molecular dynamics

The term "*ab initio*" is used mostly in the context of density functional theory, despite the fact that strictly speaking they refer to calculations based only on first principles of *Quantum Mechanics*. In most of *Ab Initio* Molecular Dynamics (AIMD), the nuclear motion is treated as classical (except of path-integral approximation). Although, direct dynamics can be coupled with a large variety of electronic structure methods and many propagation algorithms, in this thesis we mainly focus on DFT methods and the extended Lagrangian-type propagation scheme, firstly introduced by Roberto Car and Michele Parrinello in their work entitled "Unified Approach for Molecular Dynamics and Density-Functional Theory".[143] Classical mechanical treatment of nuclei and quantum mechanical of electrons are combined giving the family of techniques called AIMD. The basic idea of *ab initio* molecular dynamics methods is based on computing the forces acting on the nuclei from electronic structure calculations, that are performed “on-the-fly” during the generation of the molecular dynamics trajectory. In this way, the electronic variables are not integrated ahead, but are considered as active degrees of freedom. This means that, molecular dynamics can control complex systems and give an approximate solution of the many-electron problem. On the other hand, it implies that

the accuracy of our calculations depends strictly on how far the selected model potential is shifted to the level of selecting a particular approximation for solving the Schrödinger equation. Born-Oppenheimer molecular dynamics (BOMD), Car-Parrinello molecular dynamics (CPMD)[143] and Atom-centered Density Matrix Propagation (ADMP)[149;150] belong to this family of MD methods.

BOMD

AIMD can be carried out on the ground and excited state Born-Oppenheimer potential energy surface. The Born-Oppenheimer molecular dynamics calculations in this thesis were performed with TURBOMOLE package.[140] The electronic potential energy and its gradients are calculated quantum mechanically at the required coordinates each time step. The MD program uses the Leapfrog Verlet algorithm to turn the gradients into new atomic positions and velocities. The nuclei are propagated using classical Newtonian's equations of motion on the *ab initio* potential surface. Simulation might be performed using microcanonical (NVE) or canonical (NVT) ensembles with heating or cooling the system (thermostat for canonical dynamics) or conserving the total energy. BO approach includes the electronic structure in molecular dynamics simulations by solving the static electronic structure problem in each molecular dynamics step given by the set of fixed nuclear geometries at a given time. Therefore, the electronic structure calculation is simplified from solving a quantum time-independent problem by solving the time-independent Schrödinger equation. Finally, the electronic structure time dependency is obtained via nuclear motion. The resulting Born-Oppenheimer molecular dynamics method for the electronic ground state is defined by

$$M_A \ddot{R}_A(t) = -\nabla_A \min_{\phi_i} \{\langle \Phi_0 | \hat{H}_{eff}^{HF} | \Phi_0 \rangle\} \tag{3.17}$$

and

$$\hat{H}_{eff}(r,R)\Phi_0(r,R) = E_0 \Phi_0(r,R) \tag{3.18}$$

To obtain the Born-Oppenheimer molecular dynamics equations of motion for the electronic part i.e. Eq.3.18 lets consider special case of effective one-particle Hamiltonian such as in HF theory. Applying the variational minimum energy expectation value to the ground wave function Φ_0 defined using the Slater determinant (see Eq.2.23) and constraint that orbitals (ϕ) are orthonormal we can perform constrained minimisation of the total energy with respect to the orbitals

$$\min_{\phi_i} \left\{ \left\langle \Phi_0 \mid \hat{H}_{eff} \mid \Phi_0 \right\rangle \right\} \Bigg|_{\{\langle \phi_i | \phi_j \rangle = \delta_{ij}\}} \tag{3.19}$$

It can be inserted into Lagrange's formalism

$$\mathcal{L} = -\left\langle \Phi_0 \mid \hat{H}_{eff} \mid \Phi_0 \right\rangle + \sum_{i,j} \Lambda_{ij} (\langle \phi_i | \phi_j \rangle - \delta_{ij}) \tag{3.20}$$

where Λ_{ij} are the associated Lagrange multipliers that are necessary in order to impose the constraints. Unconstrained variation of this Lagrangian with respect to the orbitals

$$\frac{\delta \mathcal{L}}{\delta \phi_i^*} \stackrel{!}{=} 0 \tag{3.21}$$

leads to

$$H_{eff}^{HF} \phi_i = \sum_j \Lambda_{ij} \phi_j. \tag{3.22}$$

Eq.3.22 after unitary transformation of H_{eff}^{HF} leads to diagonal canonical form of Eq.2.30. Therefore, the equations of motion are define as nuclear and electronic in a HF case:

$$0 = -H_{eff}^{HF} \phi_i + \sum_j \Lambda_{ij} \phi_j \tag{3.23}$$

The main difference between BOMD and CPMD or ADMP is that the minimum of $\langle \hat{H}_{eff} \rangle$ has to be obtained at every time step of the MD simulation. Born-Oppenheimer molecular dynamics can be made as fast as other more approximated methods as Car-Parrinello molecular dynamics or Atom-centered Density Matrix Propagation methods if one sacrifice accuracy in terms of energy conservation.

CPMD

The CP scheme differs from the BO MD approach in that the wave functions are propagated together with the classical nuclear degrees of freedom using an extended Lagrangian. This based on an adaptation of the relative nuclear and electronic time-scales, which simplify the adiabatic propagation of the electronic wave function in response to the nuclear motion with appropriate large time-steps. This adjustment of time-scales through the application of a fictitious electronic wave function kinetic energy and inertia, allows the CPMD code to predict similar nuclear dynamics to the exact BO surface at smaller computational costs. Therefore, CPMD treats in a distinct way the electronic dynamics than BOMD. The CP method is essentially an extended Lagrangian dynamics scheme in which the electronic degrees of freedom do not converge iteratively at each MD step. Instead, they act as fictitious dynamical variables and are propagated along the nuclear degrees of freedom by a simple adjustment of timescales. The resultant potential energy surface stays close to the converged, real adiabatic electronic surface. Originally the CPMD code uses the Kohn Sham molecular orbitals expanded in a plane-wave basis set, and they are chosen as dynamical variables to represent the electronic degrees of freedom. The Lagrangian for the CPMD in the electronic ground state can be written as:

$$\mathcal{L}_{CPMD} = \sum_A \frac{1}{2} M_A \dot{R}_A^2 + \sum_i \frac{1}{2} \mu_i \left\langle \dot{\phi}_i \mid \dot{\phi}_i \right\rangle - \left\langle \Psi \mid \hat{H}_{el} \mid \Psi \right\rangle + constraints \tag{3.24}$$

The *dots* indicates time derivative, M are the physical ionic masses, R the nuclear coordinates, μ are arbitrary parameters of appropriate units (fictitious masses or inertia parameters assigned to the orbital degrees of freedom), $\left\langle \Psi \mid \hat{H}_{el} \mid \Psi \right\rangle$ is the potential energy of the electronic subsystem and *constraints* might be a function of both the set of orbitals $\{\phi_i\}$ and the nuclear positions $\{R_A\}$. For $\underset{\Psi_0}{min} \left\langle \Psi \mid \hat{H}_{el} \mid \Psi \right\rangle$ we are close to minimum energy, i.e. close to the exact Born-Oppenheimer surface. The equations of motion resulting for CPMD approach are then:

$$M_A \ddot{R}_A(t) = -\nabla_A \left\langle \Psi_0 \left| \hat{H}_{el} \right| \Psi_0 \right\rangle \tag{3.25}$$

$$\mu \ddot{\phi}_i(t) = -\hat{H}_{el}\phi_i + \sum_j \Lambda_{ij}\phi_j \tag{3.26}$$

The Eq.3.26 is the result of the Euler-Lagrange equation

$$\frac{d}{dt}\frac{\delta \mathcal{L}}{\delta \dot{\phi}_i^*} = \frac{\delta \mathcal{L}}{\delta \phi_i^*} \tag{3.27}$$

, where the Lagrangian derivatives are taken into account with respect to orbitals $\phi_i(r)$, where $\phi_i^* = \langle \phi_i |$.

ADMP

Schlegel et al.[141;142] proposed an alternative AIMD method: the ADMP formalism. In this method, as dynamic variables they apply the Gaussian basis sets and the one-particle density matrix ($\mathbf{P}$) within the extended Lagrangian, instead of plane-waves and Kohn-Sham molecular orbitals (χ) as in the CPMD scheme. Thus propagation of electronic variables along with the nuclei is possible through introduction of fictitious masses into each density matrix element, what leads to the simple adjustment of relative timescales. The most important advantages of using this dynamic scheme are:[151;152]

- linear scale of computational time with system size $\mathcal{O}(\mathcal{N})$;
- compact and localised wave function for a molecular system are quite well described via electronic structure calculations;
- a small number of basis functions effectively describes the state of molecular system with a desire degree of accuracy in comparison to small number of plane waves in CPMD code;

The Lagrangian for the combined nuclear-density matrix system is written as:

$$\mathcal{L}_{\mathcal{ADMP}} = \frac{1}{2} Tr\left(V^T M V\right) + \frac{1}{2}\mu Tr(\dot{P}\dot{P}) - E(R,P) - Tr\left[\Lambda\left(PP - P\right)\right] \tag{3.28}$$

P is a single-particle density matrix, M, R and V are nuclear masses, positions and velocities, respectively. The Lagrangian multiplier, Λ is introduced as a constraint on the total number of electrons, N_e and the idempotency of the density matrix. The equations of motion are defined with respect to the principle of stationary action for the Lagrangian (3.28); the propagation of the nuclei and the density in the orthonormal basis (3.15) is given by:

$$\mu \frac{d^2 P}{dt^2} = - \left[\left. \frac{\partial E(R,P)}{\partial P} \right|_R + \Lambda P + P\Lambda - \Lambda \right] \tag{3.29}$$

$$M \frac{d^2 R}{dt^2} = - \left. \frac{\partial E(R,P)}{\partial R} \right|_P \tag{3.30}$$

Chapter 4

Theoretical methods
Analysis methods.

4.1 Natural bond orbital (NBO)

Going back in history one can notice that tries to understand the nature of the chemical bond start with the development of the empirical methods and subsequent Lewis model (dot diagram of pairs of electron shared bonds and lone pairs). Lewis idea of shared electron pairs provides a qualitative picture of covalent bond and its notation is the universal foundation for chemistry world. In 1927 Heitler and London developed the first quantum mechanical description of the Lewis model. Later on Pauling, Burrau, Mulliken and Lennard-Jones contributed to a development of more complex molecular orbital description of chemical bond. The concept can be visualised according to the Aufbau principle, where the electrons delocalised over the whole molecule occupy molecular orbitals and are assigned according to the Pauli Exclusion Principle. Therefore, the set of self-consistent molecular orbitals (MO) is calculated using linear combination of atomic orbitals (AO). The number of MO is the same as the number of AO combined to form the molecule. A MO describes the behaviour of one electron in the electric field generated by the nuclei and an average field created by all other electrons occupying the other orbitals. The goal of the scientists was to invent the method that will connect the numerical content of the high quality wave function with the simple concept of the bonding theory. To understand the nature of the chemical bond between two particles A and B, one must consider two terms: kinetic energy and electrostatic forces between the nuclei and electrons, described by the quantum mechanical equations. That is achieved with the natural bond orbital (NBO) theory, which algorithms aid to accurately describe the wave function by making use of numerically optimised hybrids and pattern of the bonds, in such a localised Lewis concept. This description satisfies the accuracy of the wave function method with the simplicity of the concept to all chemists. A detailed introduction of the method and their applicability can be found e.g. here.[153;154]

Using the NBO program[155] one can get very accurate information about the

strength of the chemical bond in a many-electron molecular system in terms of localised electron-pair bonding units. The new version of NBO program (http://nbo6.chem.wisc.edu/) with optimally suited algorithms numerically solve Schrödinger's equation in the chemically intuitive way using Lewis-like bonding patterns together with 'donor-acceptor' type interactions. The program carries out the determination of natural atomic orbitals (NAOs), natural hybrid orbitals (NHOs), natural bond orbitals (NBOs), and natural localised molecular orbitals (NLMOs), and uses them to perform natural population analysis (NPA), NBO energetic analysis, and also natural resonance theory (NRT) and natural chemical shielding (NCS) analysis.

The NBO analysis makes no assumption about the wave function and use its reduced form, and hence is applicable to any general mathematical form as given by the one-electron density operator Γ:

$$\Gamma = N \int \mid \Psi(1, 2, ..., N \mid^2 dt2dt3...dtN \tag{4.1}$$

Γ can be obtained for variational, perturbative, or DFT descriptions of any type, including the exact wave function. Based on the input, the NBO bond picture is obtained from the the chosen wave function. The NBO theory relies on a method for a optimum transformation of a given wave function into a localised form, corresponding to the one-center (lone pair (LP)) and two-center (bond) elements of the Lewis like structure. To understand this transformation one it is necessary to introduce what are natural orbitals (NOs), natural atomic orbitals (NAOs) and finally natural bond orbitals (NBOs). Natural orbitals (NOs, Θ_k) are eigenfunctions of gamma, in a sense NO are the unique orbitals chosen by the wave function itself as optima for its own description. The eigenorbitals of Γ representing the natural orbitals of the wave function satisfy the equation:

$$\Gamma\Theta_k = n_k\Theta_k \tag{4.2}$$

where n_k is the electronic occupancy (population) of the eigenfunction Θ_k for the molecular electron density operator Γ of Ψ. The density operator is nothing more that the 1-electron "projection" of the full N-electron probability distribution (given by the square of the wave function $|\Psi|^2$). In this way we are able to define 1-electron subsystems of the total wave function Ψ. Therefore, we need only Ψ defining the NOs, which are its eigenorbitals, describing all single-electron properties of the wave function, including electronic density. These NOs form a complete orthonormal set The wave function Ψ is commonly described with a basis orbitals χ_k (e.g. in this thesis 6-311++G(d,p)). However, for a given Ψ the solutions of one-electron electronic density operator for NOs are in principle independent of the chosen basis orbitals. Employing basis orbitals to solve the local eigenvalue equation such as Eq. (1) is convenient from the numerical point of view. The independence of NOs (eigenorbitals Θ_k) from the chosen basis orbitals made them intrinsic and unique to Ψ, whereas basis orbitals χ_k are fitted functions, chosen to satisfy the numerical availability.

Natural Atomic Orbitals (NAOs) $\{\Theta_k^{(A)}\}$ are localised 1-center orbitals. They differ from isolated atom natural orbitals by "feeling" the molecular environment (multi-atom and multi-electron). The spatial diffuse characteristics of NAOs are optimised for the net atomic positive charge experienced by an electron in a molecular environment. They preserve intra- and inter-atomic orthogonality:

$$< \Theta_j^{(A)} | \Theta_k^{(B)} >= \delta_{j,k}\delta_{A,B} \tag{4.3}$$

Analogically to Eq.4.2, the orbitals $\{\Theta_k^{(A)}\}$ of atom A are defined as eigenorbitals of $\Gamma^{(A)}\Gamma\Theta_k^{(A)} = n_k^{(A)}\Theta_k^{(A)}$, where $n_k^{(A)}$ is the electronic occupancy (population) of the eigenfunction Θ_k for the molecular electron density operator Γ of Ψ and $n_k^{(A)} =< \Theta_k^{(A)}|\Gamma^{(A)}|\Theta_k^{(A)} >=< \Theta_k^{(A)}|\Gamma|\Theta_k^{(A)} >$. Alternatively, the $\Theta_k^{(A)}$ can be defined as 1-center "maximum population" orbitals by operations for the 1-center density operator $\Gamma^{(A)}$.

Natural Bond Orbitals (NBOs) are an orthonormal set of localised maximally occupied (mostly 1 or 2-center) orbitals that describe the Lewis-like molecular bonding pattern of electron pairs in the most accurate possible way of the total N-electron density. Similarly to NAOs, the first-order reduced electronic density operator Γ of the molecular wave function Ψ is the only input to the NBO algorithms, so the numerically determined Lewis structure representation is "natural" to Ψ itself. The NBOs are build of the so called Natural Hybrid Orbitals (NHOs) $\{h_A\}$. The NHOs also form a complete orthonormal set. Each NHO is formed from an optimised linear combination of s, p, d, etc. NAOs on the given center:

$$h_A = \Sigma_k c_k \Theta_k^{(A)} \tag{4.4}$$

for which the the overall mixing ratios can be expressed conventionally as $sp^{\lambda}d^{\mu}$ notation, where λ and μ are ratios of squared coefficients for valence p-orbitals relative to s and valence d-orbitals relative to p, respectively. Therefore, the angular composition of sp^4 hybridisation has 80% and 20% of p- and of s-character, respectively.

The type of NBOs can be bonding, lone pair, and anti-bonding. Certain delocalisation patterns arise as a weak deviation from such ideally Lewis-like localised picture and are expressed as the anti-bonding localised orbitals called non-Lewis NBOs.

The core NBOs (CR) have practically pure NAOs character. The lone pair (non-bonding) 1-center NBOs LP_A are each composed of a single normalised NHO

$$LP_A = h_A \tag{4.5}$$

The general formulation of sigma bond (σ_{AB}) between atoms A and B arises from 2-center NBOs which are normalised linear combinations of two bonding NHOs h_A, h_B. NBO uses the classical MO-type (Mulliken, Lennard-Jones) bond-orbital formulation of the electron-pair bond.

$$\sigma_{AB} = c_A h_A + c_B h_B \tag{4.6}$$

where c_A, c_B are normalised ($c_A^2+c_B^2=1$) polarisation coefficients. Its values reflect the contribution of each hybrid orbital to the bond i.e. the NBO bond polarisation may range between covalent ($c_A = c_B$) and ionic ($c_A \gg c_B$). The natural ionicity (i_{AB}) measure the bond character with values ranging from -1 (pure ionic hybrid orbital on B) through 0 (pure covalent) and 1 (pure ionic hybrid orbital on A):

$$i_{AB} = \frac{(c_A^2 - c_B^2)}{(c_A^2 + c_B^2)} \tag{4.7}$$

However, there is no unambiguous difference between 2-center σ_{AB} of highly polar form $c_A \gg c_B$ and a 1-center c_A ($c_A = 1$, $c_B = 0$). According to a fixed NBO numerical threshold highly polar form $c_A \gg c_B$ with less than 5% contribution of the different center ($|i_{AB}| > 0.95$) is define as a lone pair.

Summarising the local block eigenfunctions are obtained starting from density matrix representation of Γ in atomic orbital (AO) basis functions to localised natural atomic orbitals (NAOs) and NBOs and are natural in the sense of Löwdin, having optimal convergence properties for describing the electron density. The set of nearly double occupancy of NBOs, is said to represent the natural Lewis structure (NLS) of the molecule, which commonly describes the majority of the total electron density, as assumed in elementary bonding models. Within the pattern of bonds the chemical nature of each NBO can be identify: CR (core orbital), LP (lone pair of the valence shell with non-bonding character), LV (lone vacancy, unoccupied valence orbital), RY (unoccupied Rydberg orbital), BD (two center bonding orbital; σ, π, etc.) and BD* (two center anti-bonding orbital; $sigma^*$, π^*, etc.).

The NBO program naturally split the molecular wave function Ψ, electronic density ρ and NBO basis set $\{\Omega_i\}$ into well defined Lewis (L) and non-Lewis (NL) components:

$$\Psi = \Psi^L + \Psi^{NL} \tag{4.8}$$

$$\rho = \rho^L + \rho^{NL} \tag{4.9}$$

$$\Omega_i = \Omega_i^L + \Omega_i^{NL} \tag{4.10}$$

The NBO program also supply 3-center/2-electron bonds, 3c/4e bonds, metallic bonding and hydrogen bonding with energetic analysis of NBO interactions. Densities gotten from single-determinant MO or DFT wave functions produce Lewis and NL NBOs in a way that satisfy the one-electron eigenvalue equation of the resonance-free operator $h^{(0)}$:

$$h^{(0)}\Omega_i^{(L)} = \epsilon_i^{(L)}\Omega_i^{(L)}, i = 1, 2, ..., N \tag{4.11}$$

$$h^{(0)}\Omega_i^{(NL)} = \epsilon_j^{(NL)}\Omega_j^{(NL)}, j = N+1, N+2, ... \tag{4.12}$$

In no-resonance limit, the Lewis NBOs are occupied and NL NBOs are vacant. Due to this the concept of donors (Lewis NBOs) and acceptors (NL NBOs) of electronic occupancy arise. However, based on the application of an effective 1-electron energy

operator F (Fock or Kohn-Sham matrix) for the system the non-vanishing donor-acceptor matrix elements F_{ij} appears:

$$F_{ij} = \int \Omega_i^{(L)} F \Omega_j^{(NL)} d\tau \neq 0 \tag{4.13}$$

Estimation of the donor acceptor stabilisation energy effects is based on the second-order perturbation theory

$$\Delta E_{ij}^{(2)} = \frac{q_i |F_{ij}|^2}{(\epsilon_j^{NL}) - \epsilon_i^{(L)}} \tag{4.14}$$

, because the accuracy of the Lewis-like structures for simple molecules exceeds 99%. Details of the formula above are described in Chap.12 and in appendix F.

Among different types of analysis we use the NBO energetic analysis of the interactions corresponding to intermolecular hydrogen bonds (HB). A lot of experimental and theoretical efforts have been done in the study of this weak type of interactions.

4.2 Quantum Theory of Atoms in Molecules (QTAIM)

QTAIM: quantum theory of atoms in molecules was developed by Richard Bader and is commonly called Bader analysis.[156] The birth of the completely novel method to partition a quantum mechanical system such as a molecule is dated as 1972.[157] It was actually the future vision on how to extract the chemical information from the *ab initio* wave function.

A representation of the charge density in the real 3D space, shows that the electron density increases exponentially toward the nuclei. However if one apply a topological analysis to the electron density, many useful chemical information can be extracted. Partition of the molecule into topological atoms and use of gradient vectors, are the key concepts to get this information. The idea is to use a gradient vector field to partition full space into subspaces. The gradient vector field is a complete collection of each point so-called gradient path (GP)(trajectory of the steepest ascent through a three dimensional function), which bring itself into molecule or cluster of molecules. It can be applied to partition such 3D function as electron density (ρ), Laplacian of electron density ($\nabla^2\rho$), nuclear potential V_{nuc}, etc., leading to the concept of topological atoms. GP penetrates through constant electron density surfaces in a way to reach the local maximum. GP takes the direction orthogonal to the very small plane representing locally a surface of constant electron density value and points toward steepest upward position. The critical point (CP) is defined as a point where more than one GP meet or in other words where the density gradient vanishes ($\nabla\rho = 0$). In the 3D (rank 3) there are four types of CP: a maximum, minimum and two saddle points (point which could be a minimum or maximum depending on the direction one is scanning in). This is true

only if the curvature i.e. the function is not flat at the CP. Therefore the curvature is measured by the eigenvalues of the Hessian of the particular 3D function, electron density in our case. The Hessian is a matrix of partial second derivatives of the 3D function, where each matrix entry is defined by H_{ij}:

$$H_{ij} = \frac{\partial^2 \rho}{\partial q_i \partial q_j} \tag{4.15}$$

, where i, j=1, 2 or 3 and q_1, q_2, q_3 determine x, y and z, respectively. Therefore the matrix from of Hessian of electron density in a CP is:

$$\begin{pmatrix} \frac{\partial^2 \rho}{\partial x^2} & \frac{\partial^2 \rho}{\partial x \partial y} & \frac{\partial^2 \rho}{\partial x \partial z} \\ \frac{\partial^2 \rho}{\partial y \partial x} & \frac{\partial^2 \rho}{\partial y^2} & \frac{\partial^2 \rho}{\partial y \partial z} \\ \frac{\partial^2 \rho}{\partial z \partial x} & \frac{\partial^2 \rho}{\partial z \partial y} & \frac{\partial^2 \rho}{\partial z^2} \end{pmatrix}$$

Because this matrix is real and symmetric we can write it in a diagonal form:

$$\begin{pmatrix} \frac{\partial^2 \rho}{\partial x'^2} & 0 & 0 \\ 0 & \frac{\partial^2 \rho}{\partial y'^2} & 0 \\ 0 & 0 & \frac{\partial^2 \rho}{\partial z'^2} \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

, where the eigenvalues λ_1, λ_2, λ_3 are curvatures of ρ. The CPs are labelled with rank and signature (r, s). The rank (r) of CP denotes the non-zero eigenvalues of the Hessian and the signature (s) is the sum of the signs of the eigenvalues. According to this one can distinguish four types of rank 3 CPs: (3,-3) nuclear attractor (NA), (3,-1) bond critical point (BCP), (3,+1) ring critical point (RCP) and (3,+3) cage critical point (CCP). This classification is universal for all 3D functions so it is valid the same for the electron density as for the Laplacian of the electron density. It has to be mentioned that critical points in the Laplacian are hardly used and in any case they are not bond, ring, cage CPs, i.e. do not have the same meaning. The Laplacian of the electron density ($\nabla^2 \rho$) is second after ρ quantum mechanical function that was analysed topologically. Its values in the BCP can provide useful information of the type of the bond. Positive values of the $\nabla^2 \rho$ indicate the depletion of electron density, however the negative values ($\nabla^2 \rho < 0$) are associated with the local concentration of electron density. If the latter case is found in the internuclear region then the inter atomic interaction is classified as shared and the bond in the BCP is indicated as polar or covalent. In the opposite case, when $\nabla^2 \rho > 0$ and the number of positive eigenvalues dominates, the electron density rapidly increase away from BCP. This is a common behaviour of the electron density in the closed-shell interactions, which dominates for the ionic bond types, hydrogen bonds and van der Waals interactions. This methodology in particular focused on 3D functions like the electron density and the Laplacian of electron density were used to characterised the weakly bonded clusters of β-alanine molecules in Chap.12.

4.3 Outer Valence Green's function (OVGF)

A very useful approach to deal with the huge amount of data provided by experimental photoelectron (PE) spectra is the so-called single-particle picture (or shell model). Considering the closed-shell system as the simplest, each line in the PE spectrum can be associated with an electronic orbital energy to a first approximation. Some additional effects such as e.g. Jahn-Teller splitting gives some complications, but this approximation still provides good results of the PE spectrum and the associated individual molecular orbitals. In this sense the PE spectroscopy is strictly speaking the spectroscopy of orbitals and, via this relation it helps to understand the energetic processes occurring in ion-molecules collision (see Chap.6). Ion-molecule collisions lead to electron(s) capture/ionisation. The very first theorem related with this event is provided in 1933 as the Koopman's theorem.[158]

Koopmans theorem

The simplest method to calculate ionisation energies rely on theory provided by Koopman's theorem. This approximation seems to be proven over the decades to reasonably successful assign experimental ionisation spectra to the calculated ones. If one considers the ionic wave function Φ_i constructed from the molecular orbitals of the ground state Slater determinant analogically to Eq.2.23 without one electron taken from the spin orbital $\varphi_i(r_i)$ and leaving all others without any changes one gets the energy expression

$$E_i(\Phi_i) - E_0(\Phi_0) = I_i = -\epsilon_i \tag{4.16}$$

in which the so-called relaxation energy is neglected. The ionisation potential (the i^{th} ionisation energy) is thus approximately given by the negative value of the i^{th} energy eigenorbital. Such a statement is called the Koopmans approximation, which neglects the term responsible for charge rearrangement of the ionic wave function and also the electronic correlation energy. The total energy, the electron density and many properties are reasonably well described with the Koopmans model. Ejecting i^{th} electron by i^{th} photon out of n orbital results in a photoelectron spectrum consisting of n lines (as many lines as there are orbitals). It is because the photon interacts only with one electron at a time, where mathematically the transition operator is a one-particle operator. Thus Koopmans theorem implies that each orbital corresponds to one line in the PE spectrum. This theorem was proven to works fine for many molecules in the outer valence region of the ionisation, where two therms cancel each other. However, there is a lot of examples where the Koopmans theorem fails in determining the correct ordering of ionic states. If the energies of ionic states are quite close each other one needs a more accurate methodology to calculate the ionisation energies taking into account the effects of electron correlation and relaxation of the ionic wave function. One can get the relaxation energy from separating of SCF calculations and the

correlation energy by performing additional configuration interaction calculations for the ground and ionic states. Both methods give consistent improvement in the calculation of ionisation energies. However, there exist a direct way developed by Cederbaum,[159;160] called the Green's function method that improves the result.

Outer Valence Green Functions Method (OVGF)

The details of OVGF method can be found in the original papers[159;160] as well as in a more detailed book[161]. It might be underlined that there exists two different and distinct approximations to the one-particle Green's function method.

Lets define the one-particle Green's function in time, state space as the expectation value with respect to the exact ground state wave function of a time-ordered product of annihilation and creation operators for electrons in one-particle states

$$G_{k,l}(t,t') = -i\langle\Psi_0^N|\hat{T}a_k(t)a_l^\dagger(t')|\Psi_0^N\rangle \tag{4.17}$$

, where the $\hat{T}$ is Wick's time ordering operator, which orders the operators due to increase the time from right to left, $a_k(t)$, $\hat{H}$ is the total Hamiltonian and $a_k^\dagger(t')$ are operator in Heisenberg notation given by

$$a_k^\dagger(t) = e^{i\hat{H}t}a_k^\dagger e^{-i\hat{H}t}. \tag{4.18}$$

These operators create and annihilate electrons in one-particle states $|k\rangle$ and obey the anticommutation relation $\left[a_k, a_l^+\right]_+ = \delta_{kl}$, with all other anticommutators vanishing. A permutation of the operators from the original ordering by the action of $\hat{T}$, follows the change of sign. Applying the Fourier transform one can go over from time-state space to energy-state space

$$G_{kl}(\omega) = \int_{-\infty}^{\infty} G_{kl}(t,t')e^{i\omega(t-t')}d(t-t'). \tag{4.19}$$

The spectral representation of the Green's function can be obtained by inserting the decomposition of constant function into a sum of continuous functions. Integration leads to:

$$\begin{aligned} G_{kl}(\omega) = \lim_{\eta\to+0}\{&\sum_n \frac{\langle\Psi_0^N|a_k|\Psi_n^{N+1}\rangle\langle\Psi_n^{N+1}|a_l^\dagger|\Psi_0^N\rangle}{\omega + A_n + i\eta} \\ &+ \sum_m \frac{\langle\Psi_0^N|a_l^\dagger|\Psi_m^{N-1}\rangle\langle\Psi_m^{N-1}|a_k^\dagger|\Psi_0^N\rangle}{\omega + I_m - i\eta}\} \end{aligned} \tag{4.20}$$

, where A_n is the vertical electron affinity and I_m is the vertical ionisation energy, defined as

$$A_n = E_0^N - E_n^{N+1} \qquad \text{and} \qquad I_m = E_m^{N-1} - E_0^N. \tag{4.21}$$

So, calculating the poles of the Green's function we are obtaining both properties directly. The Dyson equation, which is formally equal to the spectral description, can be used

$$G(\omega) = G^{\circ}(\omega) + G^{\circ}(\omega)\sum(\omega)G(\omega). \tag{4.22}$$

, where $\sum_{(}\omega)$ is the so-called self-energy potential, which is the exact potential felt by an electron due to the interaction with all the others. According to the Dyson equation, the Green's function is connected via $G^{\circ}(\omega)_{kl} = \frac{\delta_{kl}}{\omega - \epsilon}$ to the HF Green's function and it posses inverses given by

$$G^{-1} = (G^{\circ})^{-1} - \sum(\omega). \tag{4.23}$$

We are interested in the perturbation expansion of $\sum(\omega)$, which is justified far away from the poles. This region is called the outer valence space, in which the satellite line structure importance is negligible. The method is named the outer valence Green's function method (OVGF). It is exact to the third order in the interelectronic interaction and with higher order terms calculated with a renormalisation procedure.[159]

Chapter 5

Theoretical methods Computational details

The Born-Oppenheimer approximation allows us to describe the essence of chemical reactivity and the behaviour of molecules in terms of a static approach - potential energy surfaces (PES) and dynamic approach - the motion on these surfaces with molecular dynamics (MD) simulations.
Any theoretical model is just a representation within the framework of the surrounding world, and as such can never be complete. The PES model is a set of approximations used to describe the small part of the real system in a particular way. The purpose is to make calculations affordable within a reasonable time and directed towards the experiments as close as possible. According to famous chemist, one of the developers of the fast Fourier transform, John W. Tukey,

> "Far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise."

For the biologically important reactions of chemical interest, we need to introduce simplifications beyond the BO approximation, i.e. into the electronic structure theories so that we can build the PES. The study of the PES helps us to understand and interpret the observed experimental spectra by calculation of the fragmentation pathways. For that purpose we used DFT and sometimes more accurate methods such as MP2 or CCSD(T) as a reference theories.

5.1 Static approach - potential energy surfaces (PESs)

Quantum chemistry calculations for the PESs study were carried out using the density functional theory (DFT) for the lowest spin state for each charge: singlet (spin equal zero) or doublet (spin equal 1/2). In particular, geometry optimisations were performed using the B3LYP functional for isolated molecules and M06-2X functional for clusters of β-alanine. Both functionals have been tested, and used in

combination with the 6-311++G(d,p) basis set. This basis contains all needed functions to properly describe positively charged amino acids and derivatives, especially it has been shown that it gives much better relative energies in comparison with the 6-31G(d,p) basis.[162;163] Moreover polarisation functions should be included for the correct description of the bonding.[164] The diffuse functions presented in a bigger basis set are used to describe properly the electronic density of ionic forms of amino acids and were applied to all conformational study in this thesis for consistency. In addition, the presence of diffuse functions together with polarisation functions has been shown to provide a better description of the PES of glycine.[164] This level of theory also has been shown to give accurate results for similar systems.[72] Harmonic vibrational frequencies have been also evaluated at the same level to characterise minima and transition states in the PES and to compute the Zero Point Energy (ZPE) correction. The absence of "imaginary" frequencies in the calculated IR spectra was interpreted as the standard condition to define the point on the potential energy surface as a minimum. On the other hand, the existence of just one "imaginary" frequency was used as a criteria of the transition state (TS) geometry achievement. For all calculated transition states, the intrinsic reaction coordinate (IRC) calculations have been also carried out to verify the minima they connect.[165] To find the TSs we used mainly the Synchronous Transit-Guided Quasi-Newton (STQN) method with two variations, QST2 and QST3. The QST2 method uses the optimised structures for the reactants and products to generate the TS. The QST3 variation requires the optimised structures for the reactants and products and also the input of a guess for the TS.[166;167] Using this methodology allowed us to find many challenging reaction transitions. Similarly to any other method for searching of TSs the one obtained with QST2 and QST3 methods were also confirmed by IRC calculations.

5.2 Dynamic approach - molecular dynamic (MD) simulations

Ab initio molecular dynamics tests were performed using S-α-alanine as a benchmark system. To this, we used several methods focusing on running MD calculations in the ground electronic state using as close as possible initial conditions. In this section we compared the methods that we decided that are somehow worth further analysis.

Benchmarking of AIMD

We present an AIMD methods benchmarking study of the fragmentation of doubly-positively-charged S-α-alanine in the gas phase. We compared the fragmentation patterns obtained with different AIMD methods presented before: ADMP, CPMD and BOMD in the singlet electronic ground state keeping as closed as possible initial conditions i.e. initial nuclear velocity distribution, fictitious mass of electrons, etc.

We run 5 trajectories for each of four initial excitation energies: 0.71, 4.27, 10.66, 14.22 eV for each of the MD method. The initial excitation energy is represented in CPMD code as initial temperature according to the formula:

$$T = \frac{2E}{fk_B} \tag{5.1}$$

, where $f = 3N - 6$ and k_B are degrees of freedom and Boltzmann constant, respectively with number of atoms, $N = 13$.

In principle the most accurate results should be obtained for a least approximated method i.e. BOMD, thus most computationally expensive. We checked also how the fragmentation patterns varies due to changing the electronic structure methods as well as different DFT-functionals. The results are presented in the Tab.5.1. Good agreement was find between ADMP and BOMD. According to the timing of the method, its simplicity and compatibility with Gaussian package, which we use for PESs studies, we decided to study isolated molecules: glycine, β-alanine, γ-aminobutyric acid with ADMP method.

Energy [K]/[eV]	CPMD(blyp)$\mu = 400$a.u.	ADMP(blyp)	ADMP(b3lyp) $\mu = 400$a.u.	ADMP(HF)	ADMP(AM1)	BOMD(b3lyp)
500/0.71	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$
	$NH_2CHCH_3^+ + COOH^+$	$NH_2CHCH_3^+ + COOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$NH_2CHCH_3^+ + COOH^+$
	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$NH_2CHCH_3^+ + COOH^+$
	$CH_3^+ + NH_2CH + COOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$NH_2CHCH_3^+ + COOH^+$
	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$
3000/4.27	$NH_2CHCH_3^+ + COOH^+$	$NH_2CHCH_3^+ + COOH^+$	$NH_2CHCH_3^+ + COOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$NH_2CHCH_3^+ + COOH^+$
	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$
	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$
	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$
	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$
7500/10.66	$NH_2CHCH_3^+ + COOH^+$	$NH_2CCH_2^+ + H_2 + COOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$NH_2CCH_2^+ + H_2 + COOH^+$	$NH_2CHCH_3^+ + COOH^+$
	$CH_3^+ + NH_2CH + COOH^+$	$CH_4 + NH_2C^+ + COOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$NH_2CHCH_3^+ + COOH^+$	$CH_3^+ + NHCH^+ + CO_2 + H_2$
	$5H + CH + CN + COOH$	$CH_3^+ + NH_2CH + COOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$NH_2CHCH_3^+ + COOH^+$	$CH_3^+ + NH_2CH + COOH^+$
	$NH_2CHCH_3^+ + COOH^+$	$NH_2CHCH_3^+ + COOH^+$	$NH_2CHCH_3^+ + COOH^+$	$NH_2CHCH_3^+ + COOH^+$	$NH_2CHCH_3^+ + COOH^+$	$NH_2CHCH_3^+ + COOH^+$
	$CH_3^+ + NH_2CH + COOH^+$	$CH_3^+ + NH_2CH + COOH^+$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3NH_2CH^+ + COOH^+$	$CH_3NH_2CH^+ + COOH^+$	$CH_3^+ + NH_2CH + COOH^+$
10000/14.22	$NH_2CHCH_3^+ + COOH^+$	$CH_3^+ + NHC + H_2 + COOH^+$	$NH_2CHCH_3^+ + COOH^+$	$CH_3^+ + NHCH^+ + C(OH)_2$	$NH_2CCH_2^+ + H_2 + COOH^+$	$NH_2CHCH_3^+ + COOH^+$
	$NH_2 + CH_2CH_2 + COOH^+$	$NH_2CCH_2^+ + H_2 + COOH^+$	$NH_2CCH_2^+ + H_2 + COOH^+$	$NH_2CCH_2^+ + H_2 + COOH^+$	$NH_2CCH_2^+ + H_2 + COOH^+$	$NH_2CCH_2^+ + H_2 + COOH^+$
	$CH_3^+ + NH_2CH_2^+ + CO_2$	$CH_3^+ + NHCH^+ + C(OH)_2$	$CH_3^+ + NH_2CHCOOH^+$	$CH_3^+ + NH_2CH + COOH^+$	$CH_4 + H + NHC + COOH$	$CH_3^+ + NH_2CH + COOH^+$
	$CH_3^+ + NH_2CH_2^+ + CO_2$	$NH_2CCH_2^+ + H_2 + COOH^+$	$CH_3NH_2CH^+ + COOH^+$	$NH_2CCH_2^+ + H_2 + COOH^+$	$NH_2CCH_2^+ + H_2 + COOH^+$	$CH_3^+ + NH_2CH + COOH^+$
	$CH_3^+ + NHCH^+ + CO + H_2O$	$NH_2CHCH_2^+ + H_2 + CO_2^+$	$NH_2CCH_2^+ + H_2 + COOH^+$	$NH_2CCH_2^+ + H_2 + COOH^+$	$NH_2CCH_3COOH^+ + H^+$	$CH_3^+ + NH_2CH + COOH^+$

Table 5.1: Fragmentation channels of Sα-alanine dication for various methods and levels of theory. We use $\Delta t_{sim} = 0.1$ fs, fictitious mass ($\mu \approx 91$ a.u., unless defined).

Convergence of ADMP method

The ADMP simulations were performed with DFT framework, in particular using the B3LYP functional with 6-31G++(d,p) basis set. Be aware of some problems with total energy conservation for MD methods[168] we carefully monitor it. It has been checked all combination between two sets of parameters such as time step (Δt =0.05, 0.09, 0.1, 0.11, 0.15, 0.2) and fictitious mass of electron (μ=0.05, 0.09, 0.1, 0.11, 0.15, 0.2) for the dication of glycine obtained after taking two outher electrons from the most stable neutral conformer of glycine.[169] In all these trajectories the simulation time was fixed for t_{sim}=200 fs and internal excitation energy settled for 0.136 eV. These tests allowed us to choose the optimal set of parameters (Δ=0.1 and μ=0.1) for the study of this kind of systems using the ADMP method. The total, kinetic electronic and kinetic nuclear energies (in eV) as a function of the simulation time (in fs) are plotted in Fig.5.1, Fig.5.2 and Fig.5.3, respectively.

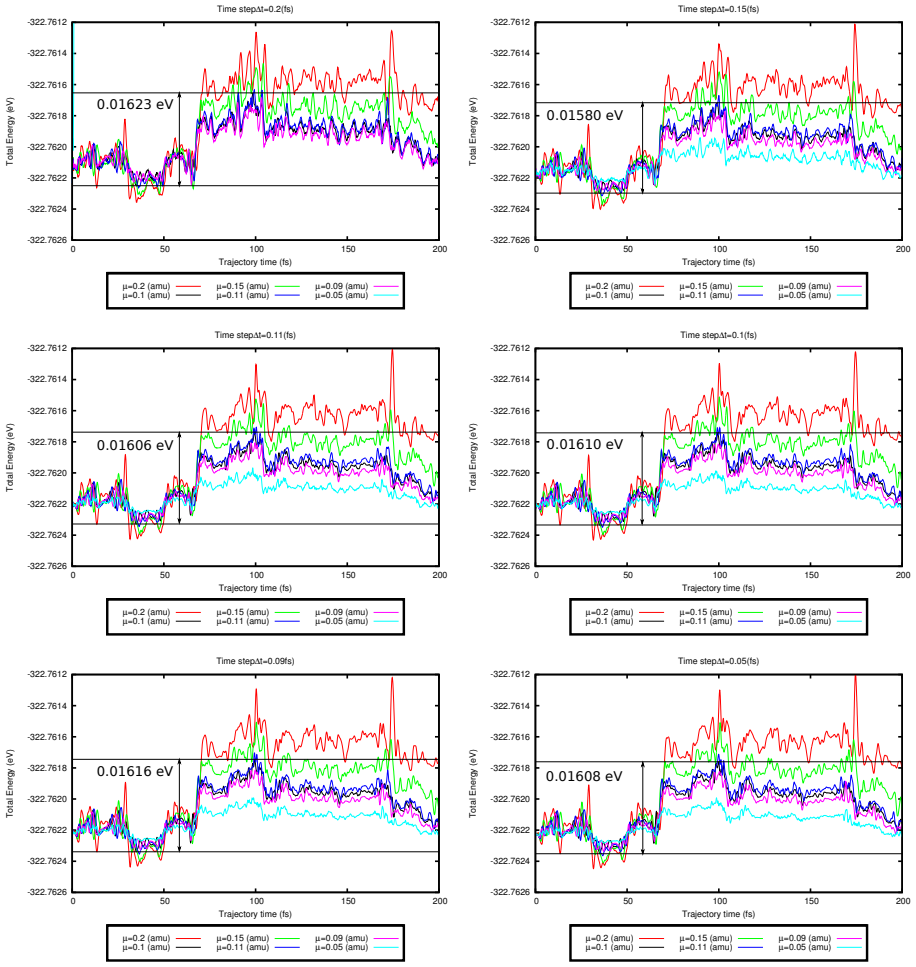


Figure 5.1: Total energy as a function of time for different time steps and fictitious mass. The black curves, parallel to x-axes represent the highest and the lowest energy value for $\mu = 0.1$ amu with different value of time step.

In all cases it is observed that the for Δt $= 0.1$ and $\mu = 0.1$ amu the conservation of the energies is acceptable, thus we do not need to reduce the time step. Finally the computational time can be saved. It is worth to mention already in this section that the changes in nuclear and electron kinetic energy increases rapidly in the first 40 fs of the simulation, suggesting the presence of a large structural changes. Indeed, these changes are due to the Coulomb explosion (CE) leading to two charged fragments.

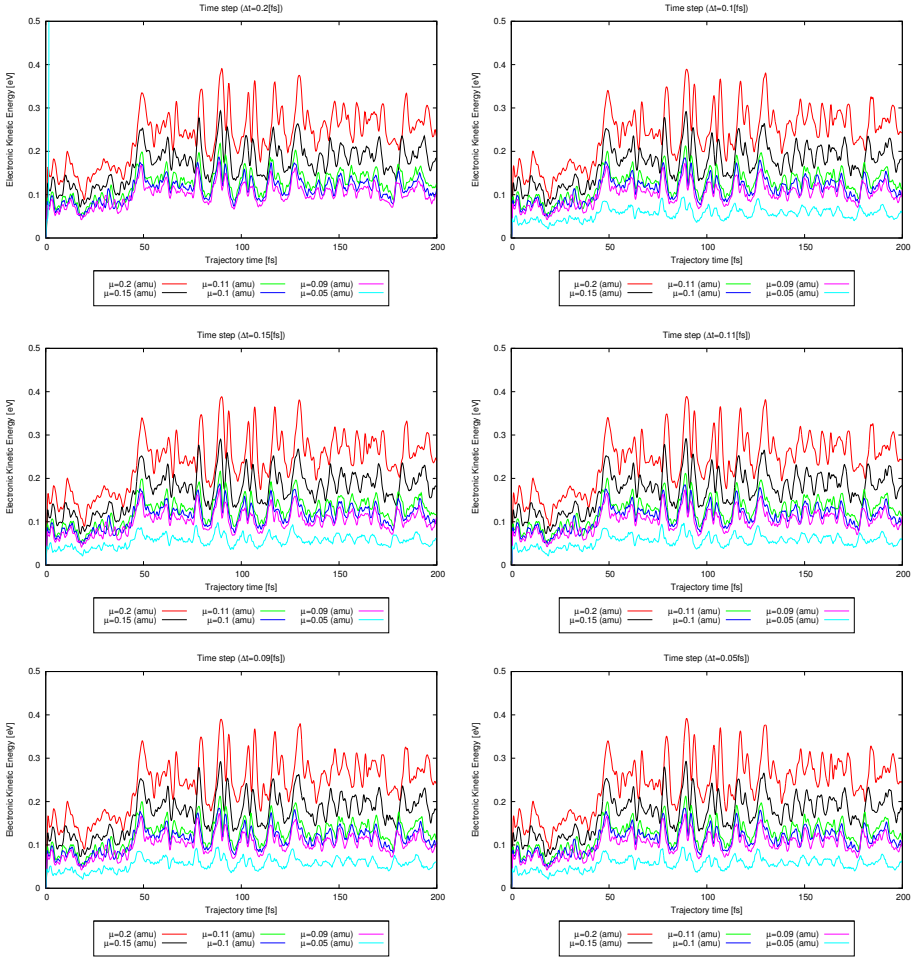


Figure 5.2: Electronic kinetic energy as a function of time for different time steps and fictitious mass.

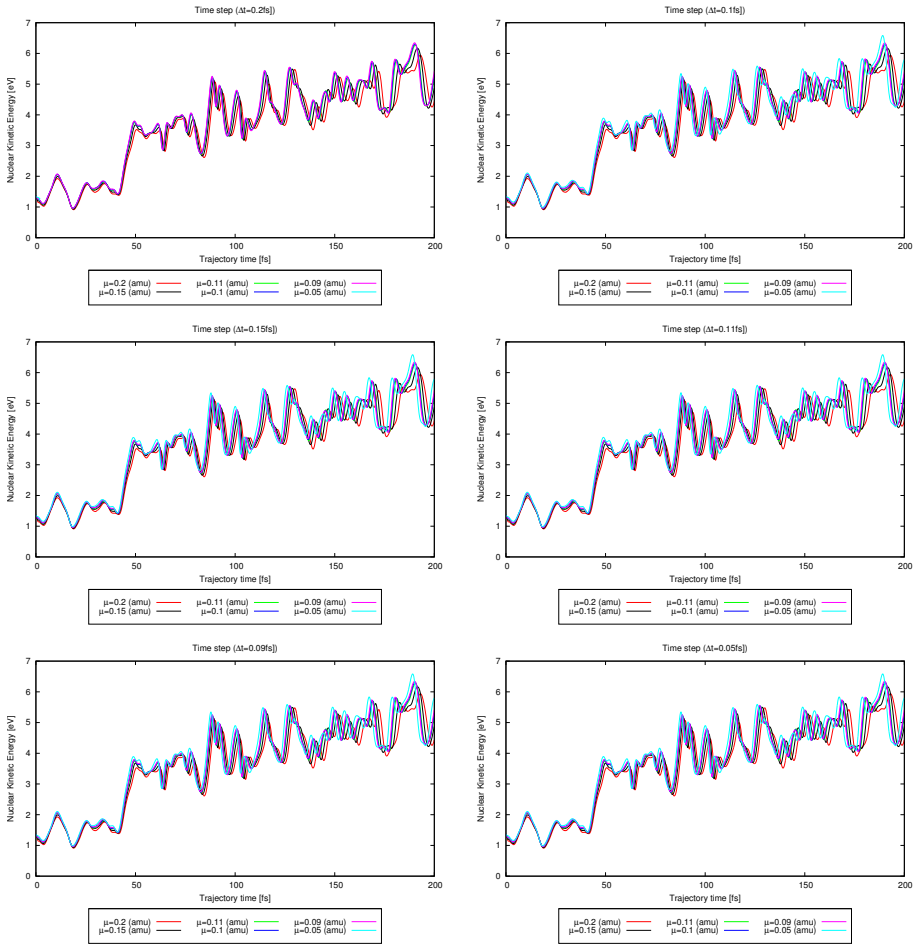


Figure 5.3: Nuclear kinetic energy as a function of time for different time steps and fictitious mass.

Part III

Results and Discussion

This part provides an overview of the results presented in this thesis. Details of each studied system, addresses the challenges of ion-molecule collision interpretation. Combination of the theoretical results with the experimental measurements is explained in the following chapters. The experimental analysis of the fragmentation of biomolecules, is carried out with mass spectrometry techniques , i.e. the intensity of positively charged molecular fragments is recorded as a function of mass over charge ratio (m/z). It focuses on the practical way to deal with interaction of molecules with highly charged ions repeated millions of time. The simplest spectrum presents the total number of counts of charged fragments as a function of m/z. Additionally, one can separate the events coming from one, two, three, etc. electron captures into 1STOP, 2STOP, 3STOP, etc. mass spectra, respectively. Moreover, the multicoincidence technique, which correlates the charged fragments of the same event is also applied. The theoretical results are obtained running molecular dynamic for a few hundreds of fs. The large numbers of trajectories are analysed to start building PES. Furthermore, many possible fragmentation pathways are calculated with barriers, including even very energetically high located fragments. Therefore, the complex concepts can be explain effectively. Indeed, calculating the potential energy surfaces (PESs), molecular dynamics (MD), molecular orbitals (MO), hydrogen bonds (HB) etc., to provide useful scientific information. The Chaps.6-10 focus on excited positively charged molecules after collisions involving highly charged ions. Adiabatic molecular dynamics in the ground state is mainly used to interpret the experimental measurements. In all these cases, the results of theoretical calculations and experimental collisions are combined to understand the stability, structure and fragmentation dynamics of these systems. Some of general features of the dynamics i.e. Coulomb explosion processes were observed together with others specific observation. The latter were further investigated using PES or wave function analysis methods. The most interesting results of these investigations are understood, using an approach based on *ab initio* atom center density matrix (ADMP) molecular dynamics simulations. Statistics were performed in a larger number of trajectories.

The results are presented in separated chapters that are based on scientific articles, they merge into a coherent "story" with completeness. Each of the chapters is a "continuation of the story" with the evolution of the system size and implementation of sophisticated methods and application of the previously understood physics to explain unexpected findings. These appear with the expansion of the complexity of system under study i.e. enlargement of the linear amino acids, nucleobasis, peptide bond prototype, derivatives of amino acid and clusters of amino acids. The first chapter in this part is mostly an experimental study where the exact distribution of the deposition energy in the ion-molecule collisions is determined. One of the main challenges in ion-molecule collisions is the determination of the energy transferred from the projectile (ion) to the target (molecule). The transfer of the energy to the molecular system drives the fragmentation dynamics. Therefore, its accurate value given as the excitation energy (above the ionisation

threshold) in the molecular dynamics simulations is very important parameter to properly mimic the experimental conditions and to get fundamental insight into ion-induced radiation damage itself. In the case of thymidine, an experimental procedure to determine the energy transfer in the ion-molecule collision was developed. Additional experiments performed with photons helped to estimate the transferred energy distribution. It fits well with the theoretical calculations based on molecular orbitals energies using the outer valence greens function method. The excitation energy is a key quantity that relates the collision with the fragmentation, and the results obtained are used in the MD calculations presented in the next chapters. We thus approach the experimental conditions as close as possible. Therefore, it makes theoretical simulations more accurate not only to interpret the experimental measurements, but also to the real biological damage caused in the framework of the hadrontherapy. Chaps.7 and 8 are based on two articles, where we focus on the stability of excited positively charged glycine after the collision with low-energy multiply charged ions in the gas phase. Firstly, we present on the fragmentation of the glycine cation, the simplest amino acids in the the lowest positive charge state. In the next chapter we present the fragmentation of the doubly-charged glycine. We show how ultrafast H migration appears in competition to well known Coulomb explosion. It is worth to indicate here that both mechanisms occur at the similar timescale ($\approx$ tens of fs). In the next chapter we elongated the amino acid chain by inserting -CH_2-, ending up with β-alanine molecule. We studied its fragmentation after extracting of two electrons with the same methodology, combining the experimental mass spectra with theoretical simulations. Again, the peaks that arise from the Coulomb explosion (CE) and from the H migration were observed. Additional peaks that can not be explained without another unexpected reaction - hydroxyl group (OH) migration were also detected. Quantum chemistry calculations allow us to predict this mechanism, as energetically the most probably.

Further insertion of -CH_2- leads to γ-aminobutyric acid (GABA). In this study we focus on the production of oxygen and nitrogen reactive species, followed by the fragmentation of dications and trications of GABA. Experimentally and theoretically we proved that after the capture two or three electrons the charge localisation is different, and thus the different types of doubly charged reactive species are produced.

In the next chapter we focused on a system with the simplest peptide bond structure - *N*-acetylglycine molecule. The single electron capture from this molecule allows us to observe microsecond timescale stable cations. Their long lifetime is possible due to the presence of the peptide bond and tautomerisation processes occurring simultaneously with charge delocalisation. This study sheds light on the additional stability against the fragmentation and atypical patterns not observed commonly in the fragmentation of isolated amino acids. It might help to analyse and understand possible ways of peptide bond formation in cluster of β-alanine after interaction with highly charged amino acids.

In the last chapter pure theoretical results of the we study of neutral clusters of β-alanine molecules are presented. Firstly we present a new computational

strategy. It consists of classical molecular simulations with AMBER-12.05 package to determine and explore widely PES. Secondly, benchmarking for the set of isomers of dimer was performed using DFT, MP2 and CCSD(T) methods. Later on geometrical end energetic criteria were chosen to evaluate stable conformers of all sizes (see Chap.12). Therefore, the extensive analysis focusing on the understanding of the production, stability, structure, dissociation and the qualitative strength of the weak interactions of the most stable structures with the best DFT functional were performed.

Chapter 6

Results and Discussion
Energy-transfer in ion-molecule collisions

The ionisation and fragmentation of the nucleoside thymidine in the gas phase has been investigated by combining ion-collision with state-selected photoionisation experiments and quantum chemistry calculations. The comparison between the mass spectra measured in both types of experiments allows us to accurately determine the distribution of the energy deposited in the ionised molecule as a result of the collision. The relation of two experimental techniques and theory shows a strong correlation between the excited states of the ionised molecule with the computed dissociation pathways, as well as with charge localisation or delocalisation.

The results presented in this chapter has been published:
S. Maclot, R. Delaunay, D. G. Piekarski, A. Domaracka, B. A. Huber, L. Adoui, F. Martín, M. Alcamí, L. Avaldi, P. Bolognesi, S. Díaz-Tendero and P. Rousseau, "Determination of Energy-Transfer Distributions in Ionising Ion-Molecule Collision", *Physical Review Letters* 117, 073201 (2016)

This chapter deals with a key aspect in the study of the fragmentation of biomolecules induced by collision with ions. It is to know what is the energy deposited in the collisional process. From the theoretical point of view this quantity is crucial because because it allows us to prepare the initial theoretical conditions for the molecular dynamics study applicable to the other biomolecules studied after collision with highly charged ions and presented in the next chapters of this thesis. Furthermore, it represents a nice example of how the combination of theory with two complementary experimental methods i.e. collision with highly-charged ion and photo-electron photo-ion coincidence (PEPICO), provides a complete picture of the fragmentation of the ionisation and fragmentation of biomolecules. The main idea is to reproduce the ion-induced fragmentation spectrum with the photon mass spectra. To correctly interpret the PEPICO experiments it is necessary to have information about the orbital energies from which the electrons are extracted and

this can only be obtained by looking at the correlated molecular orbitals and its energies, calculated with quantum chemistry methods. Later, a fitting procedure for the relevant spectra at selected excitation energies is applied, and the energy deposited on the thymidine in the ion-molecule collision is reconstructed. Most of the results presented in this chapter corresponds to the experimental work done by our collaborators but the theoretical part is critical to obtain the global picture of the process and the most relevant quantity - the deposited energy.

The understanding of the electronic and nuclear dynamics in molecular systems induced by sudden ionisation/excitation, which drives chemical reactions, offers new opportunities for controlled ultrafast chemistry.[170] For example, it has been recently observed the charge migration on the fs timescale after hole formation, which triggers atomic motion and molecular fragmentation[171] or the localisation of multiple charges in specific molecular groups and the subsequent Coulomb explosion.[172;173] Ultrafast nuclear rearrangement have been also observed in pump-probe experiments,[174] such processes compete with the expected charge separation in multiply charged molecules.[175–177] Therefore, a detailed knowledge of the response of complex molecular systems to ionisation/excitation and its influence on chemical reactivity is still today a relevant topic.[178;179] In this context, recent combined experiment/theory works have been very valuable in providing pictures of the ion-induced ionisation/fragmentation of complex molecular systems.[6;92;176;177] However, a meaningful comparison between experimental and theoretical results requires the knowledge of the energy transferred in the collision, which is in fact represented by a wide energy distribution due to interactions at different impact parameters. Pioneering experimental work reported in[180;181] has already been performed in order to determine the actual energy-deposit distributions in ion-molecule collisions, as well as to study its relationship with the observed fragmentation patterns. However, these methods require the knowledge of the initial and final projectile states which is only straightforward in the case of double-electron capture by singly charged ions (e.g. $H^+ \rightarrow H^-$), which is more the exception than the rule. In this chapter we report on the ionisation and fragmentation of a DNA building block, the nucleoside thymidine combining i) ion collisions, ii) VUV photoionisation along with iii) *ab initio* calculations. Combining such state-of-the-art techniques, we provide a complete picture of the charge localisation and the excitation energy distribution in complex molecular systems after interaction with ionising radiation. More importantly, it becomes possible to determine the energy deposited in the target as a result of an ionising collision with ions, which is the primary process associated with radiation damage. With the development of cancer therapies based on ionising particles, as hadrontherapy,[182] a better understanding of the radiation damage via a multi-scale and -disciplinary approach has become inevitable.[183] At the molecular scale, this relies on the investigation of ionisation/fragmentation of molecules of biological interest in the gas phase at different energy ranges.[6;184–186] The experiments have been performed at ARIBE, the low-energy ions facility of GANIL (Caen, France) and at the GASPHASE beamline of the synchrotron radiation facility ELETTRA (Trieste, Italy). Both experiments are

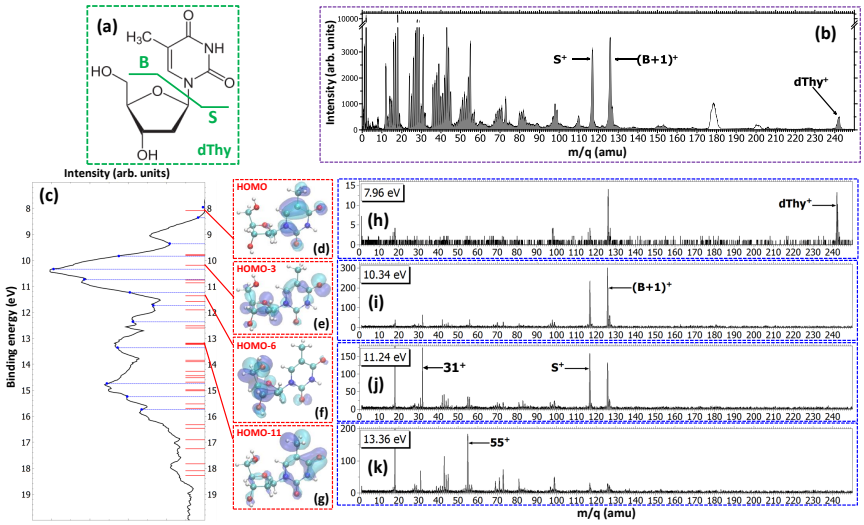


Figure 6.1: (a) Structural formula of thymidine. Considering the glycosidic bond cleavage, the fragments produced are noted **B** and **S** for the base and sugar parts, respectively. (b) Mass spectrum of thymidine after the ionisation by 48 keV O^{6+} ions. White peaks around m/z = 180 and 200 amu are due to pollutions. (c) Photoelectron spectrum (PES) of thymidine obtained at 50 eV (black curve). The blue dashed lines show the energy values chosen for PEPICO measurements. Red bars correspond to orbital energy values computed with OVGF method. Panels (d-g) show the electron density of different molecular orbitals. Panels (h-k) show the PEPICO mass spectra recorded for different binding energies of the electron corresponding to closest energies to the orbitals presented.

based on crossed-beam set-ups using coincidence time-of-flight mass spectrometry. The photoionisation experiments are based on state-selected mass spectrometry using photoelectron-photoion coincidence measurements (PEPICO). The effusive beam of neutral thymidine molecule (2'-deoxy-thymidine abbreviated dThy, see structural formula in Fig.6.1(a)) was produced by heating a powder in a oven at a temperature low enough to avoid thermal decomposition.[187] Both experimental set-ups have been described in detail elsewhere[188;189] and a brief description is given in the appendix A.

In order to have a picture of the stability of the charged thymidine in the gas phase, the mass spectrum of the charged products detected after the production of singly-charged thymidine in the interaction with 48 keV O^{6+} ions is shown in Fig.6.1(b). The peak located at m/z = 242 amu corresponds to the intact singly charged $dThy^+$ and shows that a fraction of the parent population can be stable, at least in the μs timescale, after single ionisation. The main peaks among the heaviest fragments are observed at m/z = 117 amu and m/z = 126 amu. The first one corresponds to the sugar part $\mathbf{S}^+$ whereas the second one is assigned to the

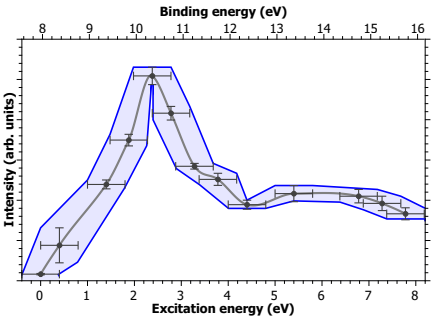


Figure 6.2: Determined distribution of the excitation energy in the ion collision (see text). The R^2 coefficient of determination for this fit is 0.86.

fragment $\mathbf{(B+1)}^+$ corresponding to an intramolecular rearrangement associated with a hydrogen transfer to the base part **B**.[190] Both fragments are the result of the glycosidic bond cleavage, an important mechanism in the radiation damage of DNA,[190–192] and contribute to 8% of the spectrum. A very small amount of fragments heavier than the base or sugar parts, i.e. loss of neutral fragments keeping intact the glycosidic bond, are also observed. This is due to the large distribution of impact parameters in the case of ion collisions[193] which leads to energy transfers, spanning from few meV to few tens of eV and involves a distribution of vibrational energy transfer and electron captures in various electronic states. Thus, the knowledge of the distribution of the energy transferred to the molecule plays a key role to unravel its fragmentation dynamics. It is difficult to assess experimentally this energy distribution even if translational spectroscopy can provide it in the case of multiple electron capture.[180;181] A method that can provide direct insight on the fragmentation following a selected energy deposition is the PEPICO technique, where the kinetic energy of the photoelectrons allows to pin point the energy left in the target.[194] The photoelectron spectrum of thymidine measured at 50 eV is shown in Fig.6.1(c). From this spectrum, thirteen photoelectron binding energy values E_b, covering the main features, have been chosen to study the evolution of the fragmentation. A simulation of the photoelectron binding energy spectra was carried out by computing the ionisation energies for the 31st highest molecular orbitals using the outer-valence Green's function (OVGF) method[160] in combination with a 6-311G(d,p) basis set of the Gaussian09 package.[195] The results are plotted in the panel (c). This method incorporates the effects of electron correlation in the computation of molecular ionisation potentials as one-particle theory for the description of ultrafast electron charge density dynamics after ionisation of an outer-valence electron. The uncertainty of the calculated energies by this method is about ± 0.3 eV.[196] The Fig.6.1(d-g) shows the computed electron densities of four orbitals corresponding to the HOMO, HOMO$-$3, HOMO$-$6 and HOMO$-$11 with binding energies of 8.09, 10.19, 11.36 and 13.18 eV, respectively, which are the closest in energy to the four selected PEPICO mass spectra shown in the Fig.6.1(h-k). They illustrate that the charge localisation after ionisation strongly depends on the orbital and may lead to different fragmentation channels. In the first mass spectrum recorded at $E_b = 7.96$ eV (Fig.6.1(h)) we observe the peak due to the parent ion, i.e. the singly charged thymidine molecule. Due to the experimental energy resolution in the PEPICO measurements, not only photoelectrons from the ground ionic state but also from deeper orbitals can be detected

(see appendix A).

This may lead to molecular dissociation and indeed some fragments are also observed. The main fragment corresponds to $\mathbf{(B+1)}^+$ indicating that the charge is mainly located on the base part as suggested by the electronic density of the HOMO (panel (d)). In the next mass spectrum, measured at E_b = 10.34 eV, two mains peaks are observed with similar intensities. They are assigned to the fragment $\mathbf{(B+1)}^+$ and $\mathbf{S}^+$ as observed in the case of the fragmentation induced by multiply charged ions. The similar intensities show that the charge has almost the same probability to be located on each one of the two moieties of the molecule as shown by the non preferential charge localisation in the orbital HOMO−3, in panel (e), and neighbouring orbitals (see appendix A). The panel (j) presents the mass spectrum recorded at E_b = 11.24 eV. The same main peaks are present, but the fragment $\mathbf{S}^+$ is now prominent. This is consistent with the preferential charge localisation of the associated orbital HOMO−6 (panel (f)), although the neighbouring orbitals can contribute to $\mathbf{S}^+$ and $\mathbf{(B+1)}^+$ peaks (see appendix A). At larger excitation energy (Fig.6.1k), the mass spectrum is characterised by a strong fragmentation showing a redistribution of the transferred energy leading to the cleavage of several bonds in the molecule.

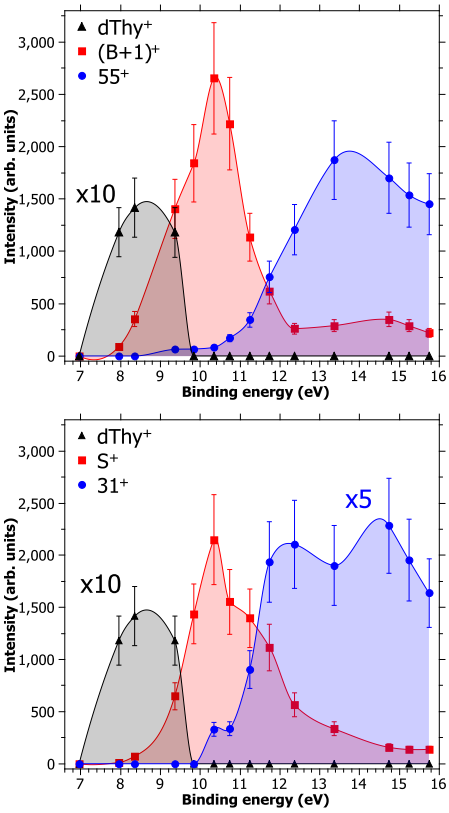


Figure 6.3: Partial ion yields of some products as a function of the binding energy obtained from PEPICO. (a) ionised thymidine ($dThy^+$) and base fragments ($\mathbf{(B+1)}^+$ and 55^+). (b) ionised thymidine and sugar fragments ($\mathbf{S}^+$ and 31^+). Each point of the curves represents the areas of a Gaussian fit to the mass spectra peaks. Error bars are estimated to 20% of the value due to the fitting method.

Using the energy selected PEPICO mass spectra we can evaluate the excitation energy distribution in an ion collision. This is achieved by fitting the results of the PEPICO spectra via a constrained linear least-square regression to the the ion-induced mass spectrum considering eleven most relevant features. The fit parameters represent the contribution of each PEPICO mass spectrum, i.e. the contribution of the fragmentation of a bunch of excited states of the singly charged ion, to the ion spectrum (see method in appendix

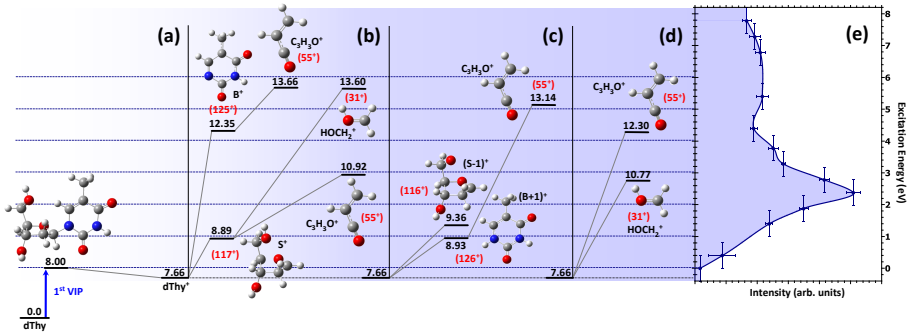


Figure 6.4: (a) First vertical ionisation potential and the most stable conformer of the singly charged thymidine ($dThy^+$). (b), (c) and (d) Energy levels of some fragmentation pathways obtained after the exploration of the potential energy surface of the most stable conformer of the singly charged thymidine. The calculated barriers of all pathways are not shown in this chapter for sake of clarity. (e) Determined distribution of excitation energy.

A). The result is displayed in Fig.6.2 as a function of the excitation energy defined as the difference between the energy left in the target and the ionisation potential. The energy distribution increases smoothly up to a maximum around $2-3$ eV and then it extends up to 8 eV and likely also above this energy, in a region not investigated in the present PEPICO experiments. Collisions at closer impact parameters can explain the extended tail towards larger deposited energy.[193;197;198] Penetrating trajectories are associated with large deposit energy of several tens of eV.[198] However, in the present interaction of 48 keV O^{6+} with thymidine, peripheral collisions leading to small energy transfer are dominating. This is due to the fact that the electron capture radius is large ($\sim$ 20 a.u. considering the classical-over-barrier model[199]) compared to the molecular size ("diameter" of $\sim$ 4.5 a.u.). The form of the distribution shown in Fig.6.2 is qualitatively similar to those obtained by fitting theoretical fragmentation probabilities to experimental measured branching ratios in small carbon clusters[193] and fullerenes,[198;200] in which the energy distribution was the fitting parameter, thus showing that the present results are also compatible with previous empirical estimations. Notice that, although the set of accessible target states can in principle be different in photoionisation and collision processes, due to the different conservation rules that can apply in each case, this is not a problem in the present work because the absence of any symmetry in the molecular target does not restrict the number of accessible states in either process. Moreover, the single-electron capture, which is the dominant process at impact energies considered in this chapter, is not accompanied by excitation of target and projectile electrons.[198;200] Therefore, one can safely assume that the mass spectra resulting from the collision involves the same target states as the PEPICO spectra.

According to the PEPICO results, the maximum in the distribution of excita-

tion energy corresponds to the region of the HOMO−3 state. The charge distribution (Fig.6.1e) leads to the cleavage of the glycosidic bond and the production of both $\mathbf{(B+1)}^+$ and $\mathbf{S}^+$ fragments as observed in ion-collision. Larger excitation energy will cause further fragmentation. The partial ion yields of the parent ion and the leading fragments in the PEPICO mass spectra are plotted as a function of the binding energy E_b in Fig.6.3. The parent ion has a maximum yield centered around 8.5 eV and then vanishes above 10 eV, while the partial yields of the main fragments $\mathbf{(B+1)}^+$ and $\mathbf{S}^+$ are observed over a wide E_b range starting around 8.5 eV and a display maximum around 10 eV (Fig.6.3). Secondary dissociation of these fragments is observed for higher E_b which corresponds to the tail towards larger excitation energies in Fig.6.2. Fragments at m/z = 55 amu have been previously assigned to $C_3H_3O^+$ arising from the base part.[190] Several pathways leading to this fragment have been calculated, as shown in Fig.6.4. The quantum chemistry calculations for the secondary fragmentation rely on an exploration of the potential energy surface in the ground state, i.e. assuming fast redistribution of the excitation energy over the vibrational degrees of freedom. The structure of the neutral molecule in the gas phase, its ionic form and the fragments produced in the relevant exit channels, together with the associated dissociation energy, have been computed with the density functional theory at the B3LYP/6-31G(d,p) level of theory, using Gaussian09.[195] The simulations show that fragment $C_3H_3O^+$ can be produced from the base part $\mathbf{B}^+$ (panel (b)) and from $\mathbf{(B+1)}^+$ (panel (c)), but also from the sugar part $\mathbf{S}^+$ (panel (b)). More surprisingly it is also possible to form this fragment directly from $dThy^+$ without glycosidic bond breaking (panel (d)). The second fragment at m/z = 31 is assigned to $HOCH_2^+$. This fragment arises from the sugar part[191] and certainly from the outside part of the furanose ring.[197] Computed formation mechanisms show that it is possible to obtain this fragment from the sugar part while keeping intact the glycosidic bond (Fig.6.4(b) and (d), respectively). Thus, combining the partial ion yields measured in the PEPICO experiments, the calculation of the binding energies of the different ionic states with the OVGF method and the dissociation pathways one can evaluate the contribution of the different fragmentation channels to the distribution of the energy transfer in the ion collision.

In summary, this chapter presents the proof-of-concept of a method to accurately determine the excitation energy distribution of complex molecular ions produced in collisions with fast ions without knowledge of the initial and final states of the projectile nor the ionisation potential of the target. The method relies on the combination of photon and ion experiments. The additional support of quantum chemistry calculations allows us to rationalise the measured energy distributions in terms of the electronic states of the singly charged ion and fragmentation channels. Thus the combination of ion and electron velocity resolved spectroscopy with *in situ* photoionisation experiments appears as a promising tool to obtain a complete picture of the molecular dynamics that follows a collision with fast ions.

Chapter 7

Results and Discussion
Stability and fragmentation of singly-charged glycine

This chapter presents a coupled experimental and theoretical work on the stability of singly charged glycine molecules ($NH_2CH_2COOH^+$, the simplest amino acid) produced in the collision of neutral glycine with low-energy multiply charged ions in the gas phase. The main fragments observed experimentally result from $C_\alpha - C_{carboxyl}$ bond cleavage. Additionally, some fragments are produced after an intramolecular rearrangement of the glycine cation, in particular the loss of a neutral water molecule following a H-migration. Moreover, this work discusses qualitatively the energy transferred to the molecule in the collision with a comparative study of the fragmentation patterns for different projectile ion charge states.

The results presented in this chapter has been published:
S. Maclot, D. G. Piekarski, R. Delaunay, A. Domaracka, A. Méry, V. Vizcaino, J-Y. Chesnel, F. Martín, M. Alcamí, B. A. Huber, L. Adoui, P. Rousseau and S. Díaz-Tendero, "Stability of the glycine cation in the gas phase after interaction with multiply charged ions", *The European Physical Journal D* 68, 149 (2014)

7.1 Introduction

With the emergence of the application of ion beams in cancer treatment (proton and hadron–therapy),[182] the study of the interaction of ionising particles with molecules of biological interest is experiencing a strong development[201–205] in order to understand the fundamental processes occurring at the molecular level.[206;207] The damage induced by the irradiation of biological tissues can be reflected in particular by the rupture of chemical bonds in biomolecular complex systems. α–amino acids are the building blocks of peptides and play several roles in the human body as for instance excitatory and inhibitory neurotransmitters.[208] The α–amino acids are

composed by two functions, a carboxyl group (COOH) and an amino group (NH_2) bonded to a central carbon, the so called carbon α (C_α). The various amino acids differ by the groups attached to the C_α as a side chain. Glycine is the simplest one, the side chain is a single hydrogen atom. Thus, its formula is NH_2CH_2COOH and its mass is 75 amu. The dissociation of positevely charged α-amino acids is mainly driven by the cleavage of the $C_\alpha - C_{carboxyl}$ bond,[45;209–211] the charge is preferentially located on the amino group due to the capture from the lone pair orbital on the nitrogen atom. This leads to the formation of the $NH_2CH_2^+$ cation ($m/z = 30$ amu).
Experimental and theoretical efforts have been joined in this thesis to study the dynamics of ionised and excited complex molecular systems in the gas phase in collaboration between the CIMAP laboratory at GANIL (Caen, France) and the Departamento de Química at the Universidad Autónoma de Madrid (Spain).[92;176;212] This chapter focuses on the fragmentation of the glycine molecule after interaction with multiply charged ions. Experimentally, coincidence measurements of cationic products analysed by mass spectrometry are used to study the fragmentation dynamics of the glycine cation. The computed dissociation pathways help to rationalise the fragmentation dynamics. Moreover, the energy deposited in the target system is qualitatively discussed by comparison of the fragmentation patterns observed in the collision with different projectiles (He^{2+}, O^{3+}, O^{6+}, Ar^{11+} and Xe^{25+}).

7.2 Methods

Experimental details

The experimental method is described in detail elsewhere,[92] thus only a short description is given here. Experiments were performed with a crossed-beam device[188;213] at the ARIBE facility, the low–energy ion beam facility of GANIL (Caen, France).[214] A pulsed beam of multiply charged ions interacts with an effusive jet of neutral glycine molecules produced by evaporation of a commercial high purity powder in an oven kept at 400 K. After the interaction, the cationic products are orthogonally extracted into a linear time–of–flight mass spectrometer. Time–of–flight spectra are recorded in an event–by–event mode allowing to determine the initial charge state of the molecular system and to measure the correlation between the charged fragments produced in a single ion–molecule collision.[92] The mass spectra for different initial charge states can be distinguished and plotted to study specific fragmentation pathways resulting from the ionisation/fragmentation of the glycine cation. In this study we focus on the fragmentation of singly charged glycine molecules thus giving rise to a single charged product.

Computational details

All calculations were performed using the density functional theory (DFT). Particularly, we used the B3LYP functional, which combines the Becke's three parameter nonlocal hybrid exchange potential[114] with the nonlocal correlation functional of Lee, Yang and Parr,[215] with the 6–311++G(d,p) basis set for exploring the potential energy surface (PES). This level of theory gives accurate results for similar systems.[72;73] Zero point energy (ZPE) correction was included. Harmonic vibrational frequencies have been calculated to characterise minima and saddle points in the potential energy surface (PES). The connections between minima (transition states) were verified by intrinsic reaction coordinate (IRC) calculations.

7.3 Results and discussion

Fig.7.1 shows mass spectra corresponding to the fragmentation of the singly charged glycine cation obtained after the interaction of neutral glycine with different multiply charged ions (Note that the experimental conditions, as beam diameter, leads to different peak width). They are constructed with the aid of the coincidence measurements (see Fig. 1.6 or Ref.[92]) and correspond to the detection of one charged particle per event in single collision mode. The heights of the peaks are presented relative to the most intense peak at m/z = 30 amu. The peak at 75 amu corresponds to the intact singly charged molecule of glycine (Gly^+ in Fig. 7.1). Thus, part of the parent molecule survives the interaction although the mass spectra are dominated by the fragmentation.

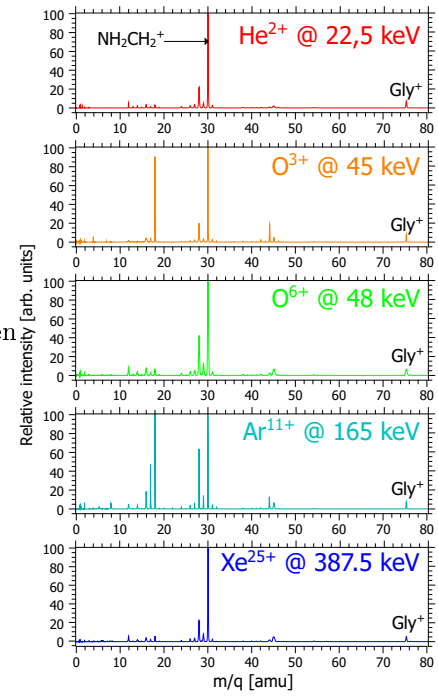


Figure 7.1: Mass spectra of the fragmentation of singly charged glycine cations obtained after the interaction of neutral glycine with different multiply charged ions He^{2+} @ 22.5 keV, O^{3+} @ 45 keV, O^{6+} @ 48 keV, Ar^{11+} @ 165 keV, Xe^{25+} @ 387.5 keV. To guide the eye, the mass spectra are normalised on the height of the peak at 30 amu. In the case of the experiments with O^{3+} and Ar^{11+} the relative intensities of the peaks at m/z = 18 and 44 amu are assigned to pollution in the residual gas. The variation of the peak widths observed for different projectiles is due to different experimental conditions.

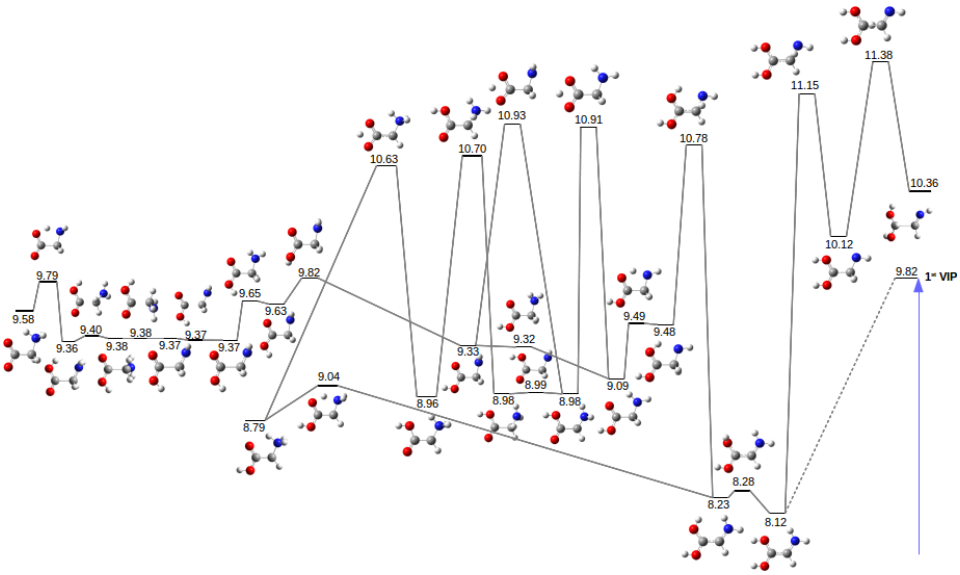


Figure 7.2: Isomerisation pathways for singly charged glycine. Relative energies, ΔE, are given in eV at the DFT–B3LYP/6–311++G(d,p) level of theory including the zero point energy correction with respect to the most stable neutral conformer of glycine.

Stability and isomerisation of Gly^+

Since part of the glycine radical cation survives the interaction, this section discusses its stability considering the isomerisation following the ionisation. Inspired by previous works Gil2008,Csaszar1992 we have chosen eight conformers of the neutral glycine molecule in its canonical form (NH_2CH_2COOH) and one geminal diol form
($NH_2CHC[OH]_2$). Experimentally the electron capture leads to a Franck-Condon transition to the ionised glycine, and thus, we optimised all these structures following a vertical ionisation and explored the potential energy surface finding the connexions between the minima obtained (see Fig. 7.2). The most stable singly charged glycine conformers show a geminal diol form. It has been shown that the isomerisation by ultrafast hydrogen transfer occurs in the case of doubly charged cation of glycine and leads to the geminal diol form.[176] The extra stability of the geminal diol was predicted in the pioneer work.[49] Indeed we observe two geminal diol forms with planar geometry and with different position of one hydrogen atom in one of the hydroxyl groups (see Fig. 7.2). The canonical form of glycine radical cation, as well as the $NH_3CHCOOH^+$ form are also predicted at relatively low energies. However, isomerisation between these structures requires energetic barriers located at more than 10 eV. This transition state is found about 1 eV higher than the entrance channel (1^{st} vertical ionisation energy).

Despite the fact that almost all transition states lie above the entrance channel, we can consider that isomerisation should be accessible as in the collision sufficient excitation energy will be transferred. However, molecular dynamics of the glycine dication have shown that the isomerisation branching ratio is about a few percent.[176] In conclusion, the isomerisation of the singly charged glycine cation explains the observation of the peak at 75 amu in the mass spectra (Gly^+ in Fig. 7.1).

Fragmentation dynamics of Gly^+

Neutral water loss

The most stable isomer of the glycine radical cation is the geminal diol form, this isomerisation following the ionisation can induced specific fragmentation pathways (see Sect. 7 and Ref.[176]). Fig.7.3 shows the presence of fragments corresponding to m/z = 57, 55 and 54 amu but with rather low intensity (<1%) which can be due to the low branching ratio towards isomerisation. The assignment of these peaks corresponds to the fragments NH_2CHCO^+, $NHCCO^+$ and $NCCOH$, respectively. Fig.7.4 shows different computed fragmentation pathways for the singly charged glycine molecule obtained after the exploration of the potential energy surface. In the left part of this figure, corresponding to the fragmentation pathways after
H-migration, one can see that the fragment NH_2CHCO^+ (m/z = 57 amu) can be easily produced without additional energy by the loss of a neutral water molecule (H_2O). This fragment can further loose a molecular hydrogen (H_2) with a quite high barrier at 13.43 eV leading to $HNCCO^+$ (m/z = 55 amu). The last step in the fragmentation path leads to $NCCO^+$ at m/z = 54 amu by the loss of H, lying around 5 eV higher than the entrance channel. As this peak is observed in the mass spectra, it is reasonable to conclude that the excitation energy of the glycine molecule after the collision can be of at least 5 eV.

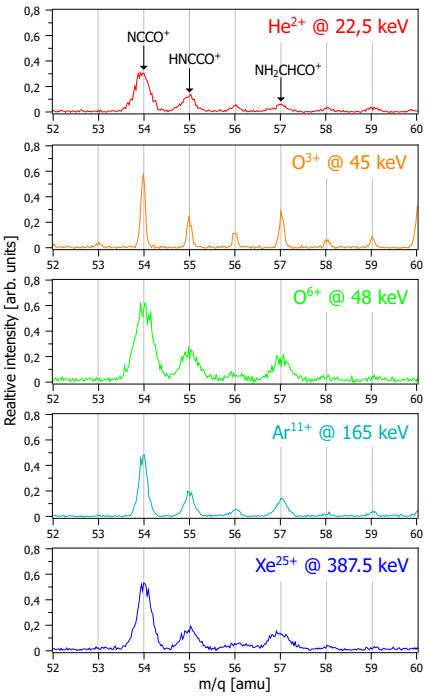


Figure 7.3: Zoom-ins in the region of interest of the mass spectra shown in Fig. 7.1. The variation of the peak widths observed for different projectiles is due to different experimental conditions.

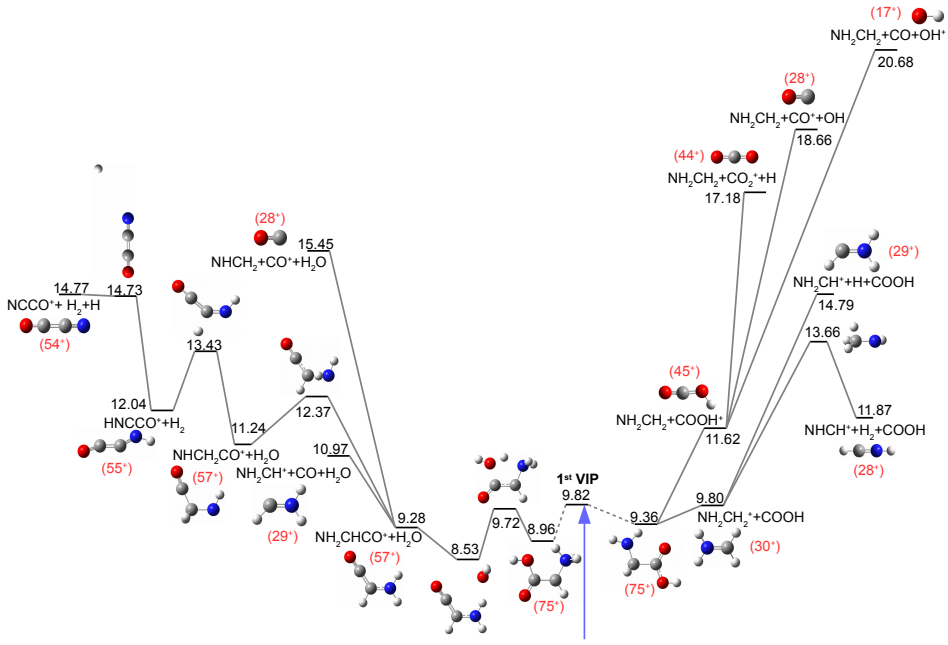


Figure 7.4: Fragmentation pathways for singly charged glycine. Relative energies, ΔE, are given in eV at the DFT–B3LYP/6–311++G(d,p) level of theory including the zero point energy correction with respect to the most stable neutral conformer of glycine. Left part corresponds to bond cleavages after isomerisation by H migration. Right part corresponds to bond cleavages on the canonical form of the molecule.

C_α–$C_{carboxyl}$ bond cleavage

The main peaks in the mass spectra are observed at m/z = 30, 28 and 45 amu (see Fig. 7.5). They correspond to the fragments $NH_2CH_2^+$, $HCNH^+$ or/and CO^+, and $COOH^+$, respectively. They are produced by the cleavage of the $C_\alpha - C_{carboxyl}$ bond. In each case the dominant fragment is $NH_2CH_2^+$ at m/z = 30 amu, the so called immonium ion (left part in Fig. 7.5). It indicates that the charge is preferentially located on the amino group. The right part of Fig. 7.4 shows the fragmentation channel producing $NH_2CH_2^+$. This channel is easily reached because it does not require excitation energy since it is found just below the entrance channel. Moreover, the minimum of energy needed to break this fragment is 3.86 eV (loss of H_2). Considering the energy distribution deposited in the molecule by the interaction with multiply charged ions (typically centred on a few eV)[217] this fragmentation channel can be reached but at a relatively high excitation energy. This can explain the predominance of $NH_2CH_2^+$ in the mass spectra. We detect

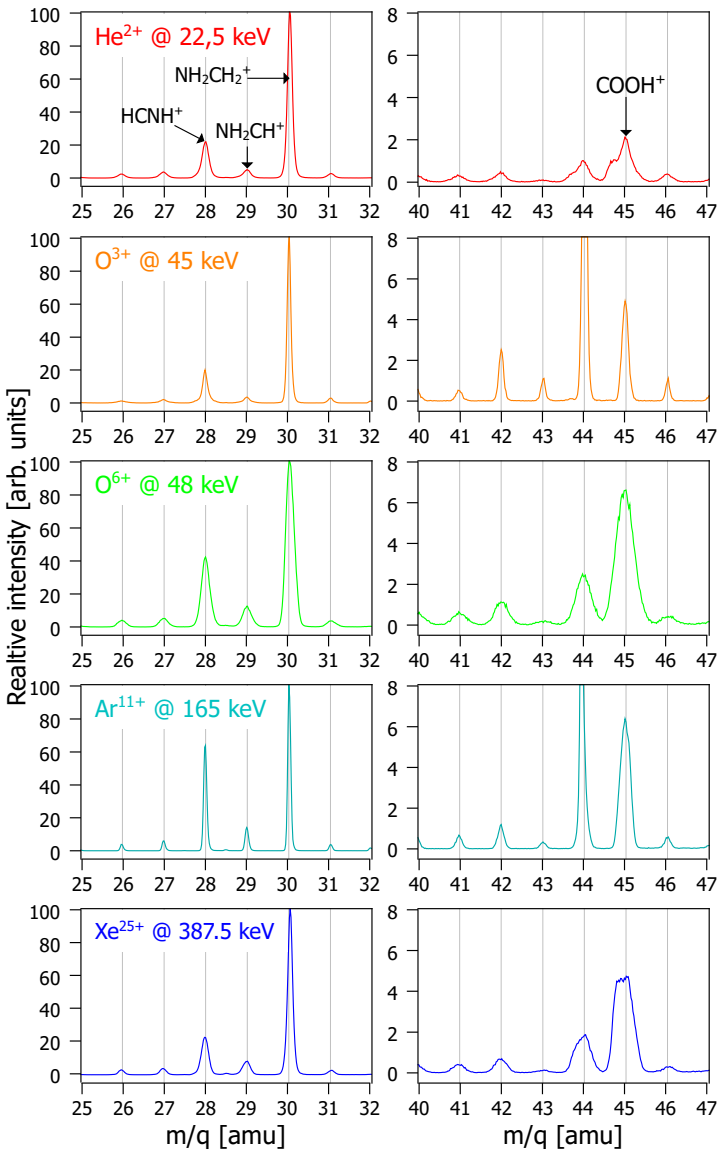


Figure 7.5: Zoom-ins in the region of interest of the mass spectra shown in Fig. 7.1. The variation of the peak widths observed for different projectiles is due to different experimental conditions.

also the fragment $COOH^+$ at m/z = 45 amu, so called carboxylic ion, leading to a different distribution of the charge in the fragmentation (right part of Fig. 7.5). This fragment is also stable due to a dissociation barrier at 5.56 eV. For these reasons these peaks are widely observed experimentally. Moreover, the ratio between the intensities of both fragments (30/45) is always higher than 4 and can increase up to 15 depending on the projectile. This is consistent with the energy required to create the immonium and carboxylic ions, as shown in Fig. 7.4, the $COOH^+$ fragment is 1.82 eV higher in energy than the $NH_2CH_2^+$ one.
The fragments at m/z = 28 amu can correspond to two different fragments $HCNH^+$ and CO^+, resulting from further dissociation of the immonium and carboxylic ions, respectively. It is reasonable to consider that both species contribute to this peak. The contribution of the fragment $HCNH^+$ comes from the fragmentation of $NH_2CH_2^+$ (Fig. 7.4 and Ref.[92]) by the loss of neutral H_2 needing 3.86 eV to be formed. The other contribution associated to the fragment CO^+ can come from two different pathways: one by the fragmentation of $COOH^+$ with 7.04 eV and one from the fragment NH_2CHCO^+ coming from the isomerisation form (see left part of Fig. 7.4) with 6.17 eV. The contribution of $HCNH^+$ could be more important since a large quantity of $NH_2CH_2^+$ ions is produced in the fragmentation yield.
Considering the fragmentation of $NH_2CH_2^+$ by the loss of a neutral hydrogen with 4.99 eV, the fragment NH_2CH^+ can be produced (m/z = 29 amu). Experimentally, the contribution of this channel is less important (around 4 times) compared to the m/z = 28 amu channel (in part because it is 1.13 eV higher in energy). But as before, an other pathway can produce this fragment from the NH_2CHCO^+ fragment (see left part of Fig. 7.4) with 1.69 eV. Although this pathway is lower in energy, it contributes less than the fragmentation of the immonium ion due to as it requires the isomerisation by H–migration of the molecule, which is statistically less probable.

Energy deposition

The distribution of the energy deposited by collisions of low-energy ions varies with the projectile charge state as the electron capture occurs at different impact parameters.[217;218] In the case of highly charged ions, the capture can occur at large distances compared to projectile in low charge states. Thus, the capture is more gentle and the excitation energy transferred to the system is reduced.

Table 7.1: Radii of single capture on the highest orbital of different projectiles.

Ion projectile	R_c (Å)
He^{2+}	3.81
O^{3+}	1.39
O^{6+}	7.38
Ar^{11+}	9.07
Xe^{25+}	14.80

To illustrate this effect, the radius of single capture R_c can be estimated with an extended classical over–barrier model (ECB)[199;219] using the computed first vertical ionisation potential of glycine. The

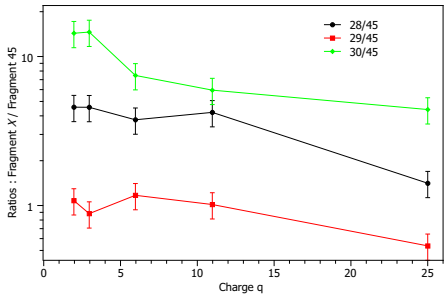


Figure 7.6: Ratios of fragments at m/z = 28, 29 and 30 with respect to the fragment m/z = 45 amu, respectively, in black, red and green as a function of the projectile charge. The error bars are estimated to 20% and are due to the fitting precision.

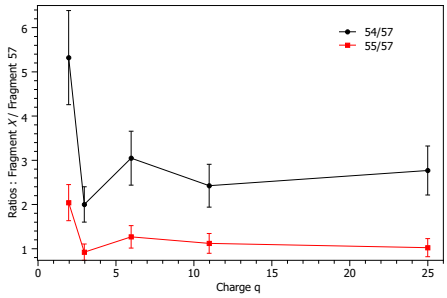


Figure 7.7: Ratios of fragments at m/z = 54 and 55 amu over 57 amu as a function of the projectile charge respectively in black and red.

radii are calculated considering that the capture occurs mainly into the highest orbital of the projectile allowed quantum mechanically. Different projectiles are studied: He^{2+}, O^{3+}, O^{6+}, Ar^{11+} and Xe^{25+}. The associated capture radii are given in Tab.7.1. As expected, the capture radius increases with the charge. However, in the case of O^{3+}, the capture radius is smaller due to the fact that some orbitals are energetically inaccessible.

In the following, the ratios of fragment intensities are obtained by a Gaussian fitting of the peaks. This allows to discuss qualitatively the stability of the fragments and the amount of energy deposited. Fig.7.6 shows the fragments at m/z = 28, 29 and 30 amu rationalised that to the fragment at m/z = 45 amu as a function of the projectile charge. These ratios decrease with the ion charge state. This means that either the amount of fragment $COOH^+$ (m/z = 45 amu) is increasing or/and the amount of fragments m/z = 28, 29 or 30 is decreasing. An explanation might be that with the increase of the charge, the $COOH^+$ species is more stable because its further fragmentation channels become less accessible as they require a high energy. Fig.7.7 shows the evolution of the ratios 54/57 and 55/57 with the charge of the projectile. A decreasing trend is observed for higher charge states. As discussed above, the distribution of the amount of energy transferred to the molecule is shifted to smaller energies with the increase of the ion charge. Thus, the fragment 57 breaks with lower probability and the fragments 54 and 55 are less produced in accordance with the computed fragmentation pathways.

7.4 Conclusion

This chapter presents the fragmentation of the singly charged glycine molecules induced by low–energy multiply charged ions in the gas phase. These experimental

results are supported by theoretical quantum chemistry calculations. These complementary approaches allow to improve the understanding of the fragmentation of the singly ionised glycine molecule. Main fragmentation channels (leading to the peaks at m/z = 28, 29, 30 and 45 amu) consist of the cleavage of the $C_{\alpha} - C_{carboxyl}$ bond. However, an intramolecular rearrangement of the ionised glycine molecule can lead to the geminal diol form explaining some of the observed cationic fragments (m/z = 54, 55 and 57 amu). By using different projectile charge states, the energy deposit has been varied. This excitation energy triggers further dissociations of the main fragments. Experimental observations of the dominant fragments can be qualitatively explained by simple energetic consideration based on theoretical values.

Chapter 8

Results and Discussion
Stability and fragmentation of doubly-charged glycine

In this chapter a combined experimental and theoretical study of the complex dynamics of excited doubly ionised glycine molecules in the gas phase is presented. Multicoincidence mass spectroscopic techniques together with ab initio molecular dynamics simulations and density functional theory calculations allow us to show that an ultrafast intra-molecular hydrogen migration (∼ 30 fs) appears in competition with the expected Coulomb repulsion.

The results presented in this chapter has been published:
S. Maclot, D. G. Piekarski, A. Domaracka, A. Meéry, V. Vizcaino, L. Adoui, F. Martín, M. Alcamí, B. A. Huber, P. Rousseau and S. Díaz-Tendero, "Dynamics of glycine dications in the gas phase: ultrafast intramolecular hydrogen migration versus Coulomb repulsion", *The Journal of Physical Chemistry Letters* 5, 3903 (2013)

After extensive study of the fragmentation of singly charged glycine induced by collision with an ions, we would like to know the behaviour of doubly charged glycine. The stability of multiply charged molecular ions in the gas phase has attracted large attention as they play an important role in gas-phase chemistry [15–18] and their stabilities challenge fundamental concepts of chemical bonding. [19;21] In general, a multiply ionised molecule becomes less stable than its neutral parent due to the extraction of electrons from the bonding regions and the charge-charge repulsion forces acting inside the molecule. The breaking of the molecule into charged fragments, often referred to as Coulomb explosion, usually leads to more stable structures. Therefore, only a limited number of molecules are thermochemically stable as dications in the gas phase.

In addition to these fundamental aspects, in the case of biomolecules, understanding the behavior of doubly charged species is also important to shed light on the radiation damage mechanisms, as swift ions can induce multiple ionisations and

subsequent fragmentations. With the advent of an ion beam based cancer treatment, the hadrontherapy[182], it is necessary to better understand the radiation damage at the molecular level.[206;220]

The detection of stable doubly charged species has also interesting consequences for astrochemistry; the enhanced reactivity shown by dications (superelectrophile effect[221;222]) has been proposed as a key feature to explain the formation of relevant new species in interstellar space[22] and planet atmospheres.[23] The search of prebiotic molecules, as glycine, in interstellar media has been a hot topic in the last decades centering much experimental[10;30;31] and theoretical efforts.[32–34]

It is thus relevant to understand how biomolecules evolve towards the stabilisation of the electrostatic repulsion forces when two or more electrons have been removed.[7;19;20] The dynamics of such multiply charged molecular species can be finely probed using physical techniques, giving insight into ultrafast chemical processes at the femtosecond (fs) timescale[223]. In this respect, the isomerisation dynamics of simple molecular cations, such as acetylene[174;224] and methanol[225], has been recently reported. It has been shown that, in such simple systems, hydrogen transfer may occur within a few tens of fs, being responsible for the ultrafast decay of the excited dications. Hydrogen transfer has also been suggested to play a role in the dissociation of protonated oligonucleotides[226] and in the formation of H_3^+ from small molecular ions.[175;211;227]

In this chapter, we present combined experimental and theoretical study to show that, in doubly charged excited glycine molecules, an ultrafast hydrogen migration process appears in competition with the expected Coulomb explosion (that leads to fission into two singly charged cations). We demonstrate that the ultrafast intramolecular hydrogen transfer is followed by a subsequent loss of neutral moieties and that this is the only plausible mechanism that explains the detection of stable dicationic species.

A similar combined experimental and theoretical strategy to investigate the fragmentation dynamics of $\gamma-$aminobutyric acid, excited and ionised in collisions with energetic multiply charged ions have been employed in the group.[92] In this thesis we improved it by including, in addition to the standard exploration of the PES, *ab initio* molecular dynamics simulations, which are crucial to obtain a complete picture of the problem. The mass spectrum of the cationic products measured after collisions of highly charged Xe^{25+} ions with glycine molecules at the energy of 387.5 keV is shown in Fig.8.1 (experimental details are given in the *appendix B*). The intact molecule survives the interaction with a certain probability, a corresponding peak is visible at $m/z = 75$ amu attributed to the $NH_2CH_2COOH^+$ radical cation. However, the spectrum is dominated by molecular dissociation with the main peaks at $m/z = 30$, 28 and 45 amu. These peaks correspond to the fragments produced by cleavage of the $C_{carboxyl} - C_\alpha$ bond leading to the formation of $NH_2CH_2^+$, $HNCH^+$ and $COOH^+$, respectively. In general, these ions are the characteristic features obtained in the fragmentation of ionised $\alpha-$amino acids.[45;209;210;228] Note that the survival yield for the intact molecule is higher after interaction with multiply charged ions compared to electron-impact ionisation or

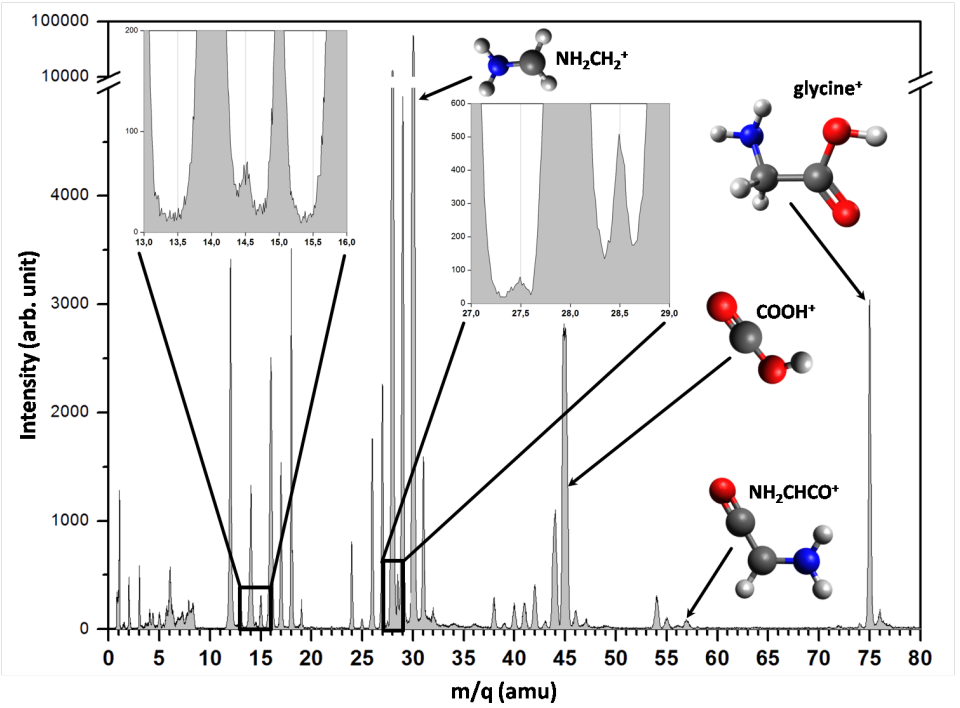


Figure 8.1: Mass spectrum of the cationic products of the interaction of Xe^{25+} ions with neutral glycine at the energy of 387.5 keV. Zoom-ins show the presence of doubly charged cations at $m/z = 14.5$, $m/z = 27.5$ and $m/z = 28.5$.

photoionisation, due to a lower energy transfer.[45] Interestingly, we clearly identify the formation of doubly charged molecular cations at $m/z = 28.5$, 27.5 and 14.5 amu (see zoom-ins in Fig.8.1). Except from the 28.5 amu peak, recently observed in electron impact experiments,[229] none of these peaks were previously detected.

A deeper analysis of the fragmentation dynamics has been carried out by means of coincidence measurements of the emitted cationic fragments from a single collision (see Fig.8.3 and details Ref.[92]). Fig.8.2 shows a so-called "correlation or coincidence map" obtained in collisions of Xe^{25+} projectiles with the glycine molecule. It displays the time-of-flight of the heavier cationic fragment (TOF2) as a function of the time-of-flight of the lighter cationic fragment (TOF1), thus reflecting the fragmentation of the glycine dication into two charged fragments. The insets in Fig.8.2 give more details in the regions of interest. The most intense islands observed in the coincidence measurements correspond to the cleavage of the $C_{carboxyl} - C_{\alpha}$ bond (island $30^+/45^+$ in Fig.8.2b) and subsequent neutral moiety emissions from the two singly charged fragments: H or H_2 from $NH_2CH_2^+$ and O or OH from $COOH^+$ (islands $30^+/29^+$, $30^+/28^+$ and $29^+/28^+$ in Fig.8.2a). Tab.B.1

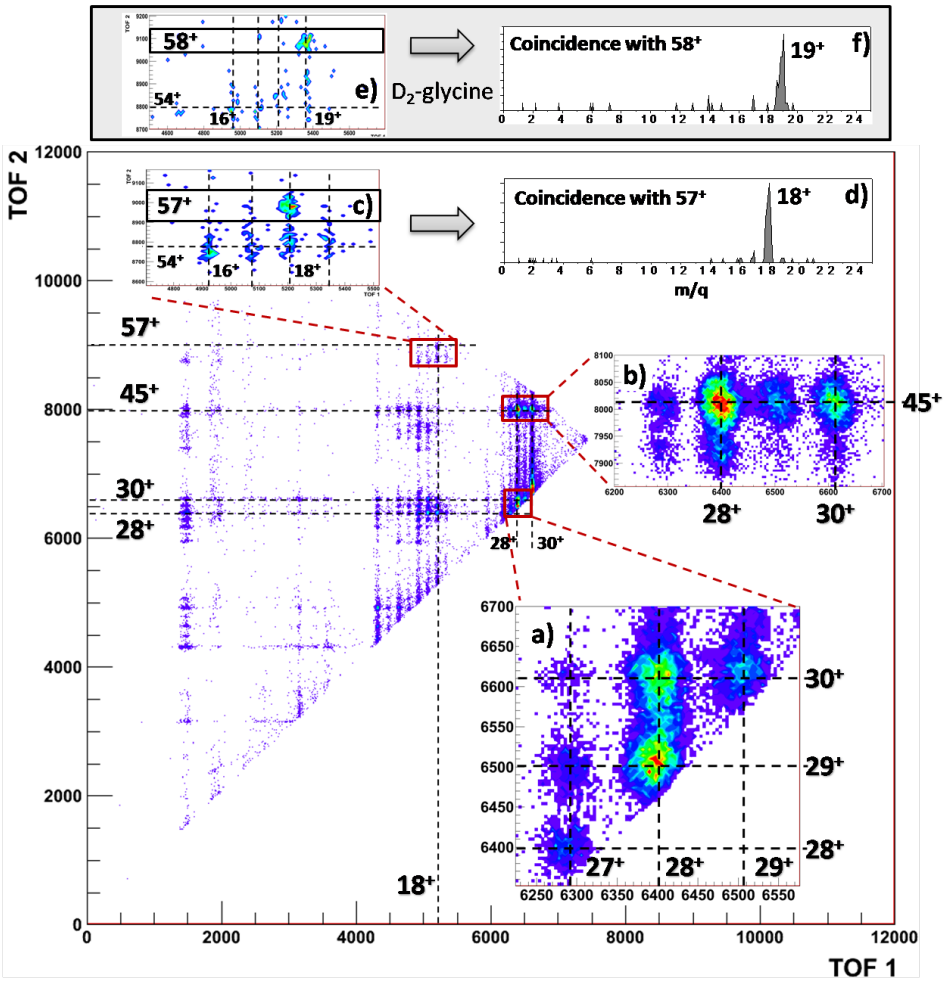


Figure 8.2: Coincidence map for the fragmentation of the glycine dication in two charged fragments after the interaction of neutral glycine with Xe^{25+} ions at the energy of 387.5 keV. The time-of-flight (in ns) of the heavier fragment (TOF2) is plotted as a function of the time-of-flight of the lighter one (TOF1). Insets show: a), b) and c) zoom-ins of the regions of interest, d) the mass spectrum of the light cationic fragments emitted in coincidence with NH_2CHCO^+, e) zoom-ins of the coincidence map for the fragmentation of the deuterated glycine NH_2CD_2COOH and f) the mass spectrum of the light fragments emitted in coincidence with NH_2CDCO^+ in collision with deuterated glycine.

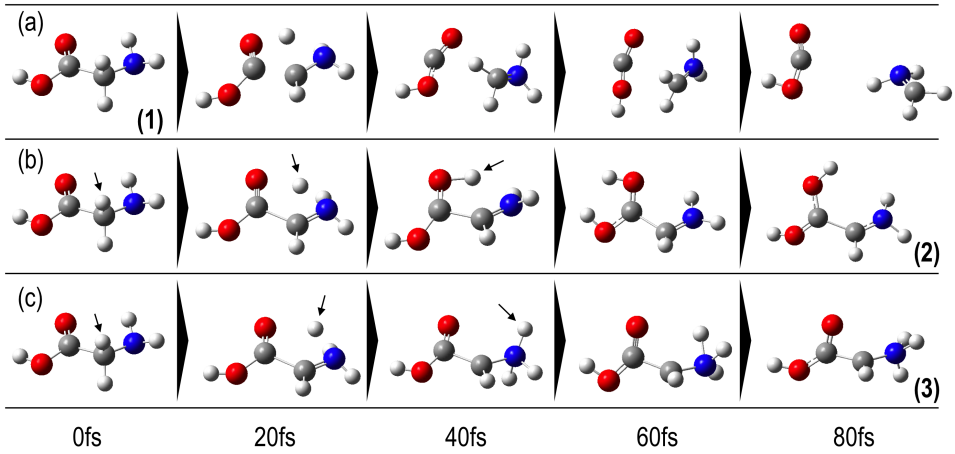


Figure 8.3: Snapshots of molecular dynamics simulations of a glycine dication (**1**) with (a) 2.18 eV of internal energy giving a fission process leading to fragments $NH_2CH_2^+/COOH^+$, b) 2.45 eV and c) 2.72 eV leading the isomerisation by H-transfer in diol $[NH_2CHC(OH)_2]^{2+}$(**2**) and $[NH_3CHCOOH]^{2+}$ (**3**), respectively. Arrows indicate the hydrogen migration.

in the appendix B gives the identification of the fragments produced with their relative intensities. As can be seen, in the coincidence map and in the corresponding "two stops" mass spectrum (see appendix B), the doubly charged species (at 14.5, 27.5 and 28.5 amu) are absent. Thus, the analysis of the coincident mass spectra shows that these doubly charged fragments are not measured in coincidence with any other charged fragment, i.e. that they are formed by neutral moiety emission from the doubly ionised glycine molecule. In addition, our coincidence measurements show also a quite interesting feature. We identify an island corresponding to $18^+/57^+$ amu (see insets c) and d) in Fig.8.2). The typical $C_{carboxyl} - C_\alpha$ bond cleavage does not explain the observation of this fragmentation channel.

The reason for these unexpected findings lies in an ultrafast intramolecular hydrogen migration, as explained by our quantum chemical calculations (see details in the appendix B). First, starting from the geometry of the most stable conformers of neutral glycine (within less than 1.5 kcal/mol with respect to the most stable one), we have carried out *ab initio* molecular dynamics simulations. We mimic the experimental conditions by extracting two electrons from the neutral conformers and introducing a certain amount of excitation energy ($E_{exc} \sim 0.1-4$ eV) randomly distributed among the internal (vibrational) degrees of freedom of the molecule. In this way, we reproduce the sudden ionisation and excitation produced in the collision of neutral gas phase glycine molecules with highly charged ions. In this respect, it is worth mentioning that, for 2-fold ionisation of Na clusters in collisions

with Xe^{25+}, the minimum energy transferred is of the order of 1-2 eV.[217] This energy range is a lower bound for the excitation energy in the glycine case, since the ion might come closer to the molecule (collision at smaller impact parameters) to capture/remove the electrons. The main fragmentation channel observed in the first steps of the dynamics leads to fission of the dicationic glycine, Gly^{2+} (**1**) into $NH_2CH_2^+/COOH^+$ (at $t \sim 30$ fs, see Fig.8.3a), which corresponds to the measured $30^+/45^+$ coincidence peak. However, the simulations also show an ultrafast intramolecular H transfer (in a similar timescale $t \sim 30$ fs) leading to two stable doubly charged isomers (see Fig.8.3b,c) $[NH_2CHC(OH)_2]^{2+}$ (**2**) and $[NH_3CHCOOH]^{2+}$ (**3**); both of them correspond to a minimum in the potential energy surface (PES) as shown further below. This occurs only starting from certain conformers and for a given range of internal energy. Coulomb explosion is the dominant channel in the energy range considered in this chapter. The importance of this channel slowly decreases with energy, especially above 1.5 eV (see appendix B for details), while that of isomerisation slightly increases. At 4 eV, the latter channel represents 5% of the observed dynamics, in reasonable agreement with the fact that H transfer channels in Fig.8.1 and Fig.8.2 are associated with peaks of very low intensity (see below). Our statistical study, performed over a total of 1700 trajectories, shows that the Coulomb explosion is indeed the dominant process. H migration is thus statistically less likely. Since both processes occur at the same timescale, the reasons for the dominant Coulomb explosion is two fold: (i) energetic -the exit channel in the Coulomb explosion is more stable than the isomer produced in the H migration- and (ii) entropic -the two fragments produced in the Coulomb explosion lead to a larger increase in entropy than the single isomer produced in the H migration.

A further theoretical analysis of fragmentation has been performed by a thorough exploration of the potential energy surface of the doubly charged glycine molecule obtained from density functional theory calculations (see Fig.8.4 and appendix B). From doubly charged glycine (**1**), the most exothermic channel is dissociation into $NH_2CH_2^+ + COOH^+$ (island $30^+/45^+$ in Fig.8.2b; see right-hand side of Fig.B.2 in appendix B). Both ions are produced with enough internal energy to follow subsequent fragmentation: $NH_2CH_2^+$ by loosing either H or H_2 (islands $45^+/29^+$ and $45^+/28^+$ in Fig.8.2b) and $COOH^+$ by loosing either OH or O (islands $30^+/28^+$ and $30^+/29^+$ in Fig.8.2a; see the four fragmentation pathways starting from $NH_2CH_2^+ + COOH^+$ in Fig.B.2 of the appendix B). Our exploration of the PES reveals that the corresponding transition states and final products are below the energy of Gly^{2+}, explaining why these are the most intense islands observed in the experiment. Note that Gly^{2+} is produced in a collision with an energetic ion, so it is reasonable to assume that the produced fragments have enough internal energy to overcome the barriers needed to reach the $45^+/28^+$, $45^+/29^+$, $30^+/28^+$ and $30^+/29^+$ fragments. Further evolution of these channels into smaller fragments is not likely because the resulting products are significantly higher in energy (see Fig.B.2 of the appendix B). Our molecular dynamics simulations show that Coulomb explosion into $NH_2CH_2^+ + COOH^+$ represents ~98% of the trajectories.

This process occurs in $\sim$ 30 fs, while further fragmentation of the singly charged $NH_2CH_2^+$ and $COOH^+$ requires longer times than those considered in the present simulations (200 fs).

As mentioned above, the system may also evolve by intra-molecular H migration keeping the two charges on the molecule. H can be transferred from C_α to produce either the geminal diol form $[NH_2 = CH - C(OH)_2]^{2+}$ (**2**) or the $[NH_3 - CH - COOH]^{2+}$ isomer (**3**) (see Fig.8.4). For singly charged glycine Gly^+, previous studies[72] have shown that the relative stability of these isomers is similar to that found in this chapter for the doubly charged ones. However, the transition states connecting them lie above the lowest fragmentation channel of Gly^+ ($NH_2CH_2^+ + COOH$); thus the isomerisation processes are not accessible channels in singly charged glycine.[37] Indeed, the mass spectra of Gly^+ is quite different from that of the singly-charged dienol isomer, since in the latter case the loss of neutral water is the dominant channel.[230;231] In contrast, the calculated PES for the doubly charged glycine shows that both isomers can be directly obtained from Gly^{2+} without energy barriers. This is confirmed by the observation of these conformers in the molecular dynamics simulations. Moreover, not only cleavage of the $C_{carboxyl} - C_\alpha$ bond, but also loss of water is observed in the experiment (see Fig.8.1). From conformers (**2**) and (**3**), at several points of the PES, competition between fission producing two singly charged species and atomic rearrangement leading to dicationic stable molecules is possible. Left-hand side in Fig.8.4 shows the most important stationary points in the PES that explain the doubly charged species detected in the mass spectrum and the $18^+/57^+$ amu island measured in the coincidence map. In more detail, from the two species produced in the H transfer, $[NH_2 = CH - C(OH)_2]^{2+}$ (**2**) and $[NH_3 - CH - COOH]^{2+}$ (**3**), we observe the formation of an intermediate complex in which H_2O is loosely bound to NH_2CHCO; from this intermediate species competition between the loss of a neutral water molecule and of a singly charged one can be seen:

$[NH_2CHC(OH)_2]^{2+}/[NH_3CHCOOH]^{2+} \rightarrow [NH_2CHCO...H_2O]^{2+} \rightarrow [NH_2CHCO]^{2+} + H_2O$

$[NH_2CHC(OH)_2]^{2+}/[NH_3CHCOOH]^{2+} \rightarrow [NH_2CHCO...H_2O]^{2+} \rightarrow [NH_2CHCO]^{+} + [H_2O]^{+}$

From the aminoketene dication obtained after evaporation of a neutral water molecule, $[NH_2 = CH - C = O]^{2+}$ (**4**), loss of CO, CO^+ and H_2 is possible:

$[NH_2CHCO]^{2+} \rightarrow [NH_2CH]^{2+} + [CO]$

$[NH_2CHCO]^{2+} \rightarrow [NH_2CH]^{+} + [CO]^{+}$

$[NH_2CHCO]^{2+} \rightarrow [HNCCO]^{2+} + H_2$

The dehydration process explains the peak at $m/z = 28.5$ amu (**4**), whereas CO and H_2 losses from the aminoketene dication explain the peaks at 14.5 amu (**5**) and $m/z = 27.5$ amu (**6**), respectively. Previous studies observing the peak at 28.5 amu[229] proposed a direct loss of OH and H from Gly^{2+}; this route would require much more energy than the two-step mechanism proposed here: isomerisation + neutral water loss. It is worth noting that (**4**) presents an energy that makes it accessible from the entrance channel (**1**) but its possible dissociation

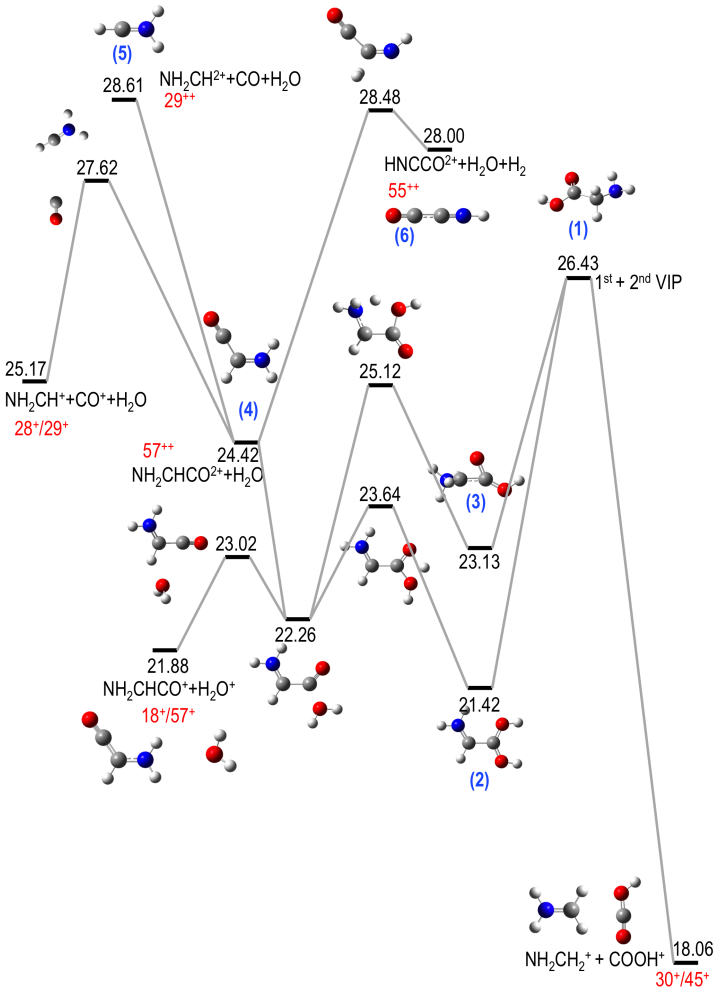


Figure 8.4: Fragmentation pathways for doubly positively charged glycine. Only the most relevant stationary points of the potential energy surface are shown. Relative energies, ΔE, are given in eV at the DFT-B3LYP/6-311++G(d,p) level of theory including the zero point energy correction with respect to the most stable isomer of the neutral molecule. Key structures are labelled as (**1**), (**2**), etc. Coincidence notation is also included, e.g. $30^+/45^+$.

pathways appear above this energy. Formation of (**6**) and (**5**) become accessible channels if we assume that part of the population of initial Gly^{2+} is formed with an internal energy of ∼2.2 eV. Thus, the computational study shows that, for the formation of the observed doubly charged species, the excited aminoketene dication $[NH_2 = CH - CO]^{2+}$ (**4**) plays a central role.

We turn now to the island at $18^+/57^+$ amu (see insets c) and d) in Fig.8.2). This is explained by the fission from the dications (**2**) and (**3**) into $NH_2CHCO^+ + H_2O^+$ (see Fig.8.4). Additional neighboring islands in the map correspond to subsequent dissociations of these cations. Again, these species can only be produced if one considers atomic rearrangement before Coulomb explosion. Thus, the detection of the $18^+/57^+$ island is an additional experimental evidence of the intramolecular H transfer. To further confirm this mechanism, we have performed experiments with the doubly deuterated glycine molecule in the central carbon atom (NH_2CD_2COOH). In this case, as expected, we observe the coincidence NH_2CDCO^+/HDO^+ island at $19^+/58^+$ amu (see insets e) and f) in Fig.8.2). Although in the fragmentation of singly charged enol glycine it was shown that dehydration involves the amine group,[49] our study of the potential energy surface (Fig.8.4) predicts the formation of the complex $[NH_2CHCOH...H_2O]^{2+}$ after H migration from (**2**) and from (**3**) prior to the fission into $NH_2CHCO^+ + H_2O^+$.

Finally, we want to stress that methylenimine (CH_2NH)[232] and cyanoformaldehyde (CNCHO)[233] have been detected as neutral species in star-forming regions; they are isomers of the doubly charged species NH_2CH^{2+} and $HNCCO^{2+}$ detected in the present work. Reaction of these dicationic species with neutral CO and H_2, respectively, can lead to the formation of H_2NCHCO^{2+} (**4**), which in turn can react with water to form the diol isomer of Gly^{2+} (**2**). These reactions are exothermic and the energy involved in the process would be enough to obtain Gly^{2+} (**1**). Although, according to our results, alternative routes may also lead to the formation of $H_2NCHCO^+ + H_2O^+$ or $HNCCO^+ + H_2O^+ + H_2$, the mechanism proposed here suggests that this path might lead to the spontaneous synthesis of glycine in the interstellar media.

In summary, we have studied the fragmentation dynamics and different deexcitation pathways characterising the decay of glycine dications in the gas phase. Experimentally, excited molecular dications are produced in collisions with Xe^{25+} ions and the fragments are identified by coincident time-of-flight mass spectrometry. We find that two types of processes occur in competition. On the one hand, charge separating processes leading to the coincident detection of two singly charged fragments. On the other hand, the formation of smaller stable dications produced by evaporation of neutral fragments following ultrafast intramolecular hydrogen migration (∼ 30 fs). Only a few per cent (typically 5-10%) of the detected processes correspond to H-migration, in agreement with the *ab initio* molecular dynamics simulations in which the internal energy of the molecule is larger than ∼ 2 eV. These processes are expected to occur in other biomolecules and therefore should be considered when trying to obtain a complete picture of their complex fragmentation dynamics.

Chapter 9

Results and Discussion
Stability and fragmentation of doubly-charged β-alanine

This chapter presents a combined experimental and theoretical study of the fragmentation of doubly positively charged β-alanine molecules in the gas phase. The dissociation of the produced dicationic molecules, induced by low-energy ion collisions, is analysed by coincidence mass spectrometric techniques; the coupling with ab initio *molecular dynamics simulations allows rationalisation of the experimental observations. The present strategy gives deeper insights into the chemical mechanisms of multiply charged amino acids in the gas phase. In the case of the $\beta-$alanine dication, in addition to the expected Coulomb explosion and hydrogen migration processes, we have found evidence of hydroxyl-group migration, which leads to unusual fragmentation products, such as hydroxymethyl cation, and is necessary to explain some of the observed dominant channels.*

The results presented in this chapter has been published:
D. G. Piekarski, R. Delaunay, S. Maclot, L. Adoui, F. Martín, M. Alcamí, B. A. Huber, P. Rousseau, A. Domaracka and S. Díaz-Tendero, "Unusual hydroxyl migration in the fragmentation of β-alanine dication in the gas phase", *Physical Chemistry Chemical Physics* 17, 16767, (2015).

9.1 Introduction

Dicationic species, extensively studied due to their unusual stability and reactivity,[221;222] are involved in the creation of new biomolecules in planets atmospheres[23] and in the interstellar media.[22] Numerous amino acids were detected in Carbonaceous Ivuna chondrites, a type of meteorite, with the simplest β-amino acid, β-alanine ($NH_2CH_2CH_2COOH$) as the most abundant one.[10] This finding can be considered as the manifestation of spontaneous gas phase synthesis of amino acids

in the interstellar space.

Moreover, previous theoretical and experimental works showed the important role played by organic radical cations in various chemical reactions,[58;72–74] where the first step is the ionisation of the molecule. For instance an important electrophilic modification of amino acids which takes place after ionisation or nitrosation of alanine[234–236] produces lactic acid.[237] Therefore, there is a need to understand the chemical behaviour of charged amino acids.

In particular, the fragmentation of positively charged β-alanine has been the object of extensive investigations.[169;211;238;239] In single ionisation of β-alanine, and other small amino acids, it was found that the loss of the carboxylic group is one of the dominant processes.[169;211;238] In the α-alanine case, photoionisation experiments showed that the principal cationic fragment produced corresponds to $NH_2CH_3CH^+$, formed by the loss of the COOH radical in a simple $C_{carboxylic} - C_\alpha$ bond cleavage.[210] $C_\alpha - C_\beta$ and $C_\beta - C_\gamma$ bond breaking are the dominant channels for β-amino acids[228] and γ-amino acids[92], respectively. For more complex amino acids side chain bond breaking appears in competition.[169;238] For double ionisation, simple Coulomb explosion is expected to dominate along the same bond cleavage mentioned above.[92;176]

In this chapter, we also join theoretical and experimental efforts. *Ab initio* molecular dynamics and exploration of the potential energy surface (PES) together with multiple-coincidence mass spectrometry techniques allow us to identify all possible exit channels and fragmentation mechanisms following the ionisation processes for excited doubly charged β-alanine in the gas phase. We have found unusual fragmentation mechanism implying hydroxyl group (OH) migration. They appear in competition with other processes such as hydrogen transfer[176] or Coulomb explosion. These findings are of high relevance for a better understanding of radiation damage processes, since hydroxyl radicals have been identified as one of the main factors responsible for such damage.[240–243] Indeed, it has been shown that in biological tissues the damages produced in the biomolecules are not only caused directly by the particle-matter collision but also by radicals and secondary particles created along the ionisation path.[4;50] Radiation damage of biological tissues at the molecular level has been widely studied by employing experimental techniques involving interaction of biomolecules with ionising particles (X-rays, electrons, ions) in the gas phase[206;207;244;245] in combination with theoretical simulations.[92;176;182;206;220]

Moreover, the migration of the hydroxyl group has been shown to be a crucial step in the mechanistic processes of enzymes, such as the propanediol dehydratase,[246;247] and has been widely evaluated by theoretical simulations.[248–250] The proposed processes in these former studies differ from the mechanisms presented in this chapter, mainly because in doubly-charged $\beta-$alanine, the charge drives the first steps in the fragmentation and thus, the OH migration can be considered as a consequence of the bond breaking due to Coulomb explosion. The importance of OH migration is highlighted here because it leads to ionic products detected with high intensity in the experiments ($NH_2CH_2^+$ in coincidence with CH_2OH^+), *i.e.* they come from a dominant channel.

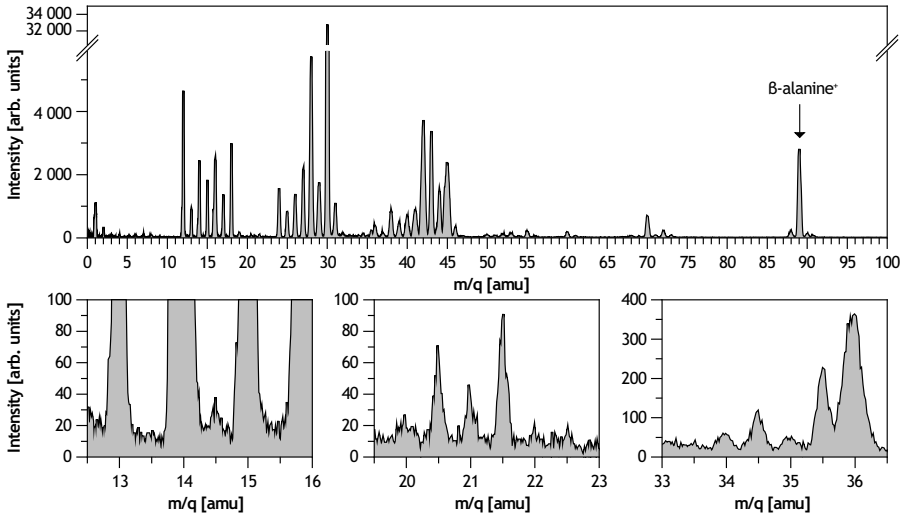


Figure 9.1: Mass spectrum of cationic products from collisions between O^{6+} ions at 48 keV and neutral β-alanine molecules in the gas phase. 3 zoom-ins enlight the presence of doubly positively charged products in the mass range of 13-16, 20-23 and 33-36 amu.

9.2 Methods

Computational Details

Quantum chemistry calculations were carried out using the density functional theory (DFT). In particular, geometry optimisations were performed using the B3LYP functional, which combines the Becke's three parameter non-local hybrid exchange potential[114] with the non-local correlation functional of Lee, Yang and Parr.[215] This functional has been used in combination with the 6-311++G(d,p) basis set. This method has shown to provide an accurate description of the PES of cationic amino acids.[72;92;164;176;238] Harmonic vibrational frequencies have been also evaluated at the same level to characterise minima and transition states in the PES and to compute the Zero Point Energy (ZPE) correction. For the obtained transition states, intrinsic reaction coordinate (IRC) calculations have been carried out to verify the minima they connect.[165]

Ab initio molecular dynamics (AIMD) were performed using the Atom-centered Density Matrix Propagation method (ADMP),[141;142;149;150] employing the B3LYP functional together with the 6-31++G(d,p) basis set. The simulations were carried out using a time step of $\Delta t = 0.1$ fs, a fictitious mass of $\mu = 0.1$ amu and a maximum simulation time of $t_{max} = 300$ fs. To mimic the experimental conditions we have performed the AIMD simulations in the gas phase by introducing a certain amount of excitation energy, $E_{exc} \approx 0.03-4.08$ eV, and extracting the two electrons from the highest occupied molecular orbital (HOMO) in a Franck-Condon type

transition. This energy range covers the estimated excitation energy in collisions of highly-charged ions with amino acids in the gas phase[92;176;251]. In the collisions of other type of biomolecules, such as adenine with singly charged ions, a higher excitation energy has been measured ($\sim$ 8 eV) due to the different conditions.[181] The internal energy was randomly distributed over all the vibrational degrees of freedom in each trajectory. In this way, we reproduce the sudden ionisation and excitation process produced in gas phase collisions of ions with neutral β-alanine molecules. The AIMD simulations were carried out considering the 12 most stable conformers of neutral β-alanine in the gas phase.[55] For each isomer and each value of the excitation energy we run 20 trajectories. Statistics were then performed over these trajectories. All the calculations were performed using the Gaussian09 program.[195]

Experimental details

Products of the interaction between O^{6+} ions and isolated β-alanine molecules in the gas phase were studied by means of coincidence time-of-flight (TOF) mass spectrometry. The experiments were performed at the low energy ion beam facility ARIBE of GANIL in Caen, France. The experimental setup has been described in detail elsewhere,[188] only a brief description is given here. The O^{6+} ions beam was extracted from the electron cyclotron resonance ion source, mass selected, pulsed, collimated and transported in the interaction zone where the beam crossed a molecular effusive jet of neutral β-alanine molecules produced by the evaporation of a powder from a heated oven device at 400 K. Cationic products of the interaction were extracted into Wiley-McLaren linear time-of-flight mass spectrometer.[213] At the end of the free-flight region of the spectrometer (TOF tube of 1 m length) these products were accelerated and impacted a gold-coated plate producing secondary electrons which are detected by a micro-channel plate detector after deflection by a weak magnetic field produced by Helmholtz coils. This Daly-type detector allowed efficient and uniform detection of the interaction products. The arrival times were measured in an event-by-event mode over a wide mass-over charge (m/z) ratio range. Each event was characterised by the number of detected fragments and their associated time-of-flight. This allowed to construct the correlation between the charged fragments proceeding from a single ion-molecule interaction (for details see Ref.[92]).

9.3 Results and discussion

The total mass spectrum of positively charged products from collisions of O^{6+} at 48 keV ions with neutral β-alanine molecules in the gas phase is shown in Fig.9.1. A strong peak associated with the intact molecule (m/z=89 amu) is observed and represents 3% of the total intensity. However, the mass spectrum is dominated by fragmentation products. The dominant exit channels for heavier fragments

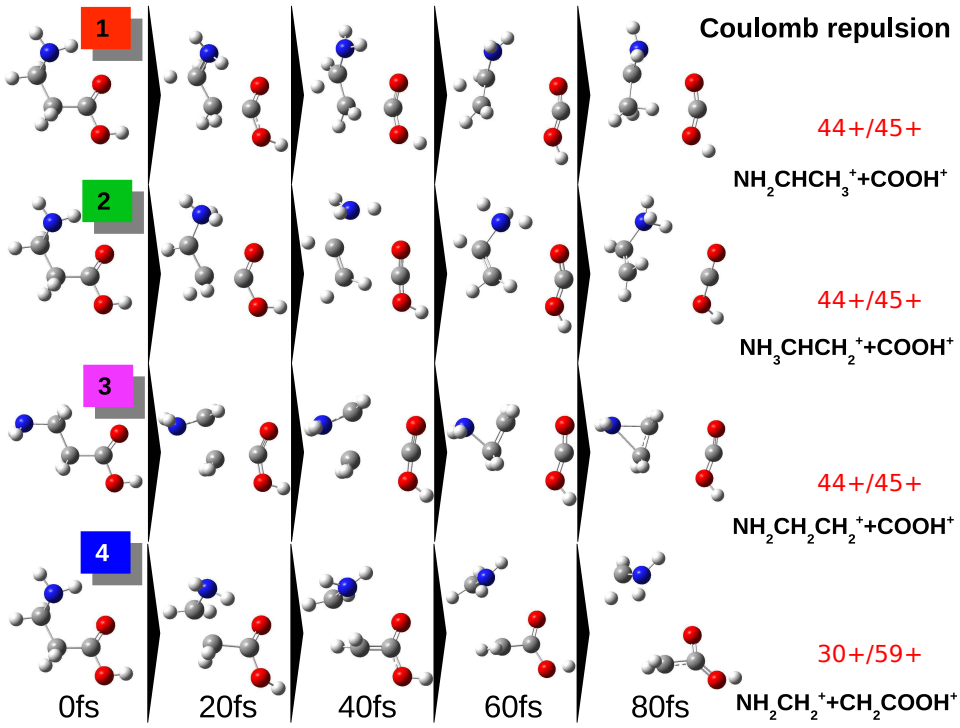


Figure 9.2: Snapshots of molecular dynamics simulations of a β-alanine dication showing Coulomb explosion, *i.e.* charge separation processes: a) $C_{carboxylic}$−C_α bond breaking: (1) $NH_2CHCH_3^+/COOH^+$, (2) $NH_3CHCH_2^+/COOH^+$, (3) $NH_2CH_2CH_2^+/COOH^+$; b) $C_\alpha - C_\beta$ bond breaking: (4) $NH_2CH_2^+/CH_2COOH^+$.

(m/z≥24 amu) come from $C_{carboxylic} - C_\alpha$ and $C_\alpha - C_\beta$ bond cleavages and subsequent fragmentations. The peak at m/z=30 amu dominates the spectrum and corresponds to the iminium cation ($NH_2CH_2^+$). Fragment ion assignments to the most intense peaks have already been made in previous works.[210;211] Interestingly, in the inclusive mass spectrum we can also observe the presence of molecular dications like in the case of the simplest amino acid glycine.[176] The different chemical mechanisms leading to cationic fragments are inferred with the quantum chemistry simulations. We first present the theoretical results obtained in the *ab initio* molecular dynamics simulations. In the 4080 trajectories performed for the doubly charged β-alanine, we have characterised seventeen different mechanisms that can be classified into three groups: (i) Coulomb explosion, (ii) isomerisation and (iii) combined processes. Snapshots of theses processes are shown in Figs.9.2, 9.3 and 9.4, respectively. Trajectories 1, 2, 3, 4, 5, 9 and 10 are observed with larger probability and represent the main fragmentation mechanisms. Detailed information of all mechanisms is given in the appendix C.

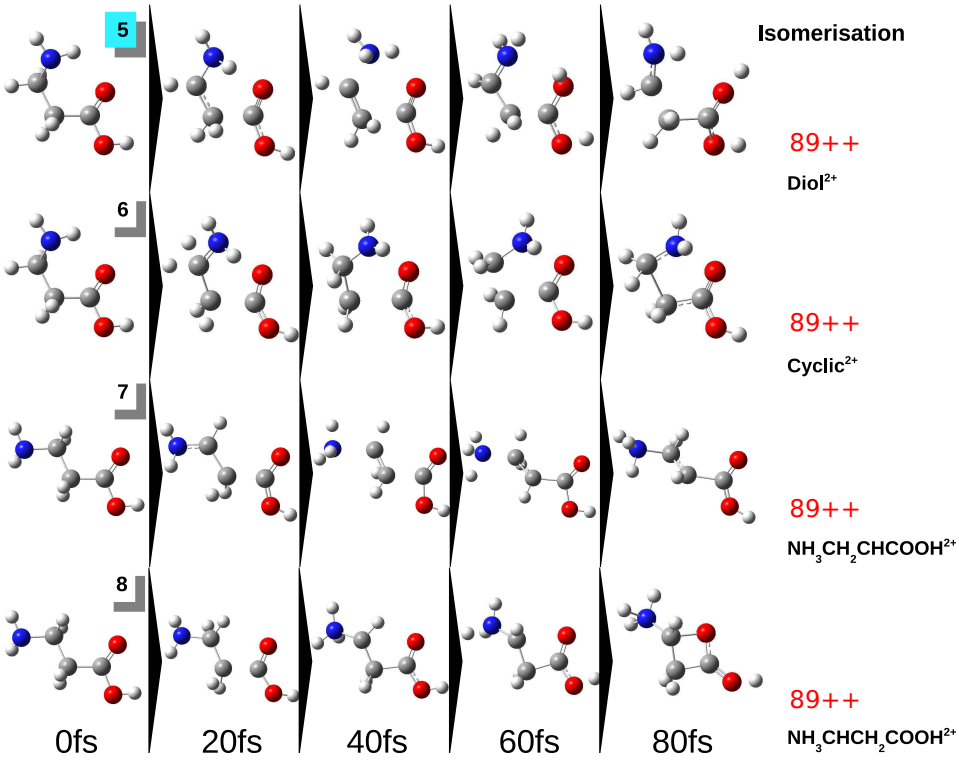


Figure 9.3: Snapshots of molecular dynamics simulations of a β-alanine dication showing isomerisation processes: (5) diol^{2+}, (6) 5 – membered ring dication, (7) linear $NH_3CH_3CHCOOH^{2+}$ and (8) 4 – membered ring dication.

The Coulomb explosion dominates (see Fig.9.5a). Four different Coulomb explosion processes can be distinguished (see Fig.9.2): $C_{carboxylic} - C_\alpha$ bond cleavage leading to (1) $NH_2CHCH_3^+ + COOH^+$, (2) $NH_3CHCH_2^+ + COOH^+$ and (3) $NH_2CH_2CH_2^+ + COOH^+$; and $C_\alpha - C_\beta$ bond cleavage leading to (4) $NH_2CH_2^+ + CH_2COOH^+$. Among them, channels (1) and (4) strongly dominate with ≈40% and 30%, respectively (see Fig.9.5b). In the processes (1) and (2) we observe H migration to the terminal C and N atoms, stabilising the produced iminium and ammonium cation, respectively. On the other hand, the mechanism (3) is an example of geometry reorganisation, leading to a stable cyclic isomer of the $NH_2CH_2CH_2^+$ cation. The carboxyl ion ($COOH^+$) produced in (1), (2) and (3) is very stable. However, fragmentation of this ion was evaluated for other organic systems leading to strongly observed signals at 28^+ and 29^+, associated with CO^+ and COH^+, respectively.[252] Mechanism (4) leads to the very stable $NH_2CH_2^+$ cation, which dominates the mass spectrum.

Fig.9.3 shows the time evolution of different isomerisation processes of β-alanine

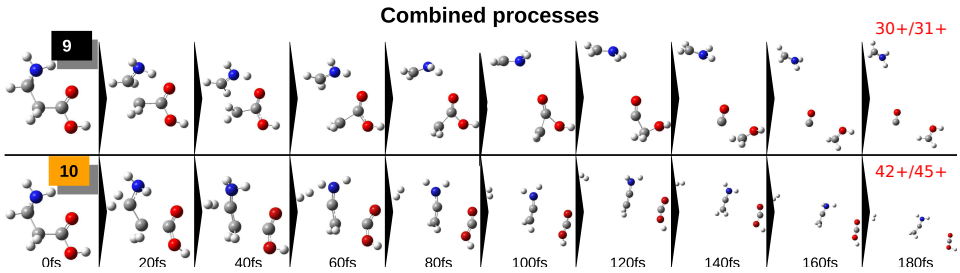


Figure 9.4: Snapshots of molecular dynamics simulations of a β-alanine dication showing combined processes: (9) OH migration, (10) Coulomb explosion and hydrogen molecule emission; all combined mechanisms available in the appendix C.

leading to the diol^{2+} (5), 5-membered ring^{2+} (6), linear $NH_3CH_2CHCOOH^{2+}$ (7) and 4-membered ring^{2+} (8) structures. Only one of them does not show H transfer and leads to the stable cyclic form of β-alanine^{2+} (mechanism (6) in Fig.9.3). This structure has not been previously reported, but a similar 5-membered ring intermediate (ozalone) was proposed by Harrison,[253] after fragmentation of a peptide followed by further elimination of water. 5-membered ring structures were previously observed after hydrogen migration and water emission in the homologous series of carboxylic acids $HOOC(CH_2)_nCOOH$.[252] In competition with cyclisation the formation of linear $NH_3CH_2CHCOOH^{2+}$ is also observed, which requires H transfer and appears sporadically in the simulations. H transfer in amino acids is commonly observed in solution[236;254] and has been recently reported in the gas phase after the interaction with highly charged ions.[176]

Combined processes are defined as those trajectories that show Coulomb explosion and molecular rearrangement simultaneously. They are also associated with the emission of neutral moieties. The most important and dominant complex molecular rearrangement of those observed (see Fig.9.4) is the migration of the hydroxyl group, leading to the emission of neutral carbon monoxide: $NH_2CH_2^+ + CH_2OH^+ + CO$ (mechanism (9)). This mechanism appears in competition with Coulomb explosion showing $\approx$ 15% of probability. Another observed channel is $NH_2CCH_2^+ + H_2 + COOH^+$ (10). Mechanisms leading to emission of neutral CO_2, NH_3 and CH_2 also appear (see the appendix C) but they are very rare processes (with less than 2 % of probability).

It is important to notice that the fragmentation pattern strongly depends on the geometry of the conformers (see details in the appendices). This can be explained by the fact that the structure of the dication formed at the initial stage of the fragmentation (in the ionisation) determines the reaction path and, thus, the fragmentation products.[255] However, since the temperature in the oven is 400K, only the most stable neutral conformers will be populated before the collision and thus, not all the conformers included in the calculations will contribute to the fragmentation pattern because they are too high in energy to be present in the beam.

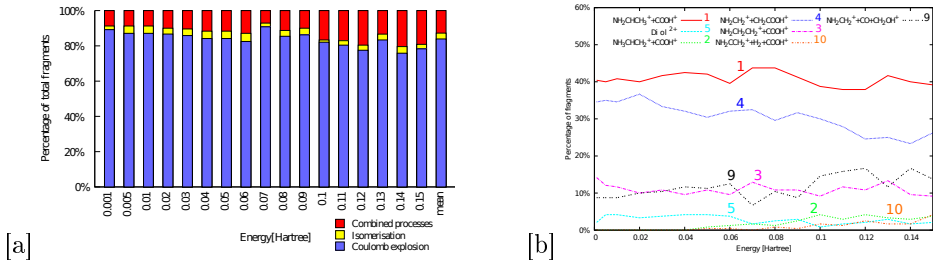


Figure 9.5: *Ab initio* molecular dynamics statistics for all neutral conformers of β-alanine and percentage of the most important fragmentation channels. Both quantities are given as a function of the internal excitation energy (in Hartrees). A detailed analysis for the statistics of each neutral conformer is included in the appendix C.[a]Competition between Coulomb explosion vs. isomerisation vs. combined processes for all studied conformers of β-alanine with relative energy $\approx$ 2 kcal·mol^{-1}. [b]Percentage of the most important fragmentation channels as a function of the excitation energy (in Hartree) for all studied conformers of β-alanine with relative energy $\approx$ 2 kcal·mol^{-1}.

Statistical details about all mechanisms are presented in the appendices.

Molecular dynamics simulations were performed up to 300 fs. To evaluate the energetics and the mechanisms of the subsequent processes we have thoroughly explored the potential energy surface (PES). The exploration of the PES has been performed starting from the most probable and relevant processes as observed in the statistics of the molecular dynamics simulations: *i.e.* mechanisms (1), (2), (3), (4), (5), (9) and (10). Firstly, we calculated all of the possible pathways corresponding to the mechanisms obtained in the molecular dynamics simulations. Later, we associated the lowest energy channels with the peaks observed in the experimental measurements. Using this approach we obtained a complete picture of the fragmentation dynamics and further energetic information of each channel which becomes comparable with the experimental measurements, done at the μs timescale. We present these results separately considering the following order: (a) $C_{carboxyl} - C_\alpha$ bond cleavage, (b) $C_\alpha - C_\beta$ bond cleavage, (c) H migration and (d) unexpected OH migration.

(a) $C_{carboxyl} - C_\alpha$ bond breaking.

Fig.9.6 shows zooms in the most interesting regions of the correlation map, which displays the time-of-flights (in ns) of the heavier fragments (TOF 2) as a function of the time-of-flights of the lighter ones (TOF 1). It shows the islands with the highest intensities with special emphasis on the ion pairs ($42^+/45^+$), ($30^+/42^+$), ($30^+/31^+$) and ($28^+/30^+$). Direct bond cleavage of β-alanine does not explain most of these islands. The assignments of such species and the understanding of their production require some complementary studies. Quantum chemistry calculations allow us to propose the mechanisms underlying the formation of the different molecular

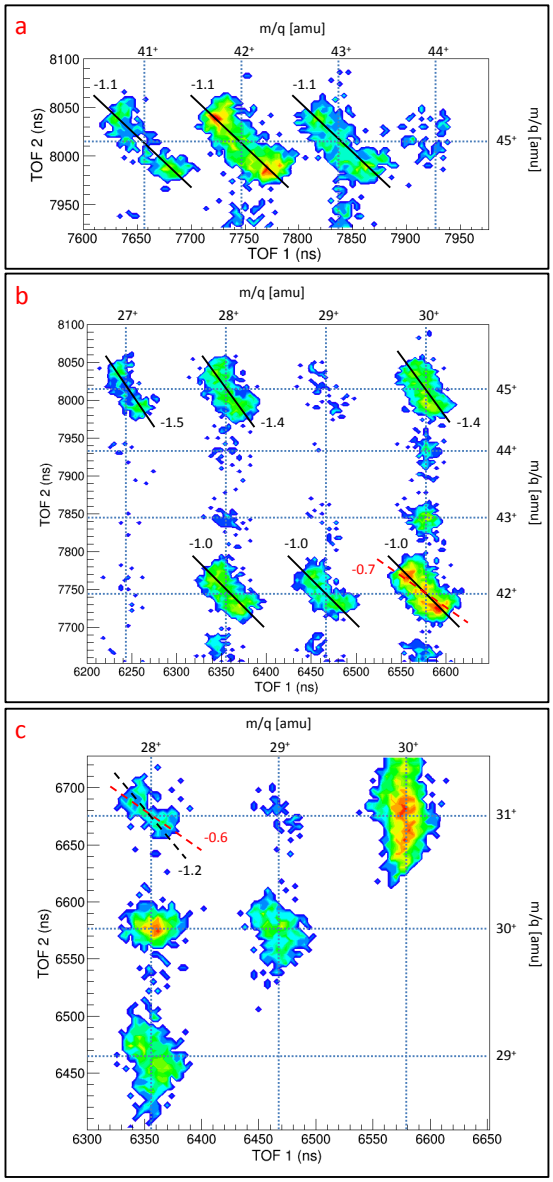


Figure 9.6: Coincidence map for the fragmentation of the β-alanine dication into two charged fragments after the interaction of neutral β-alanine with O^{6+} ions at the energy of 48 keV. Regions of interest correspond to islands: a) 45^+ with other heavy fragments (43^+, 42^+, 41^+), b) 42^+ and 45^+ with the NH_xCH_x fragments ($1 \leq x \leq 2$), c) lighter fragments coming from OH migration and associated with the fragment 31^+. Experimental slopes values are associated with solid lines and dashed lines indicate different computed mechanisms.

cations observed after $C_{carboxyl}-C_\alpha$. Starting from the $NH_2CHCH_3^+ + COOH^+$ exit channel we can easily explain the peaks observed in the zoom on the coincidence map (Fig.9.6a). Fig.9.7 presents the part of potential energy surface corresponding to these pathways. The most intense islands related to $C_{carboxyl} - C_\alpha$ are those corresponding to $COOH^+$ (m/z=45 amu) in correlation with other species: m/z=44, 43, 42 and 41 amu. Processes corresponding to further fragmentation of the $COOH^+$ appear at much higher energy and thus, the corresponding islands contribute in much lower percentage in the correlation map: slight signals coming from $COOH^+$ fragmentation due to O or OH emissions are observed, similarly to those previously reported in the literature.[252] The coincidence measurement at m/z=($44^+/45^+$) amu presents low intensity. This is in agreement with the simulations assuming that $NH_2CHCH_3^+$ (m/z=44 amu) possesses enough internal energy to undergo further fragmentation since this channel is almost 9 eV below the entrance channel ($1^{st} + 2^{nd}$ vertical ionisation potential, VIP). The observation of the islands at ($43^+/45^+$) and ($42^+/45^+$) amu corresponds to the emission of H and H_2 from this fragment. The strong intensity of these islands in the correlation map (Fig.9.6a) confirm the proposed mechanisms. The peak at m/z=43 amu is assigned to $NH_2CCH_3^+$ and $NH_2CHCH_2^+$ (see Fig.9.7); they are cationic forms of acetaldimine and vinylamine, respectively. After emission of one hydrogen atom, dehydrogenated acetaldimine ($NHCCH_3^+$) and vinylamine ($NH_2CCH_2^+$) cations (m/z=42 amu) become more stable than their original forms. The intramolecular hydrogen transfer between them is an exact analogue to the previously studied imine/enamine tautomerisation processes.[256]

The analysis of the slopes in the detected islands provide information of the fragmentation mechanisms[257] (see appendices for details): the pair ($43^+/45^+$) corresponds to the fragments $NH_2CHCH_2^+$ or $NH_2CCH_3^+$ with $COOH^+$. A slope of -1.02 is expected considering this mechanism. Quantum chemistry calculations are thus in a good agreement with the measured slope. For the pairs ($42^+/45^+$) and ($41^+/45^+$), in which two or three hydrogen atoms are emitted from $NH_2CHCH_3^+$ (m/z=44 amu), we obtained the slopes -1.05 (-2H) and -1.07 (-3H) with the computed mechanisms. They are also in agreement with the experimental ones taking into account the experimental uncertainties. The conclusion is that computed pathways are consistent with the experimental measurements (see Tab.C.1 in appendix C).

Complementary information on the fragmentation of $NH_2CHCH_3^+$ is given in the right part of Fig.9.8. The pathway after CH_2 emission leaves $NH_2CH_2^+$, a very stable ion in coincidence with $COOH^+$ ($30^+/45^+$). Further emission of hydrogen produces $NH_2CH^+/COOH^+$ and $NHCH^+/COOH^+$, where the latter channel (direct H_2 emission) is much lower in energy than the H emission (which appears 1.3 eV above the entrance channel). Accordingly, only the one corresponding to hydrogen molecule emission is observed in the experiment, pair ($28^+/45^+$) (see Fig.9.6b).

Other fragments observed in the correlation with $COOH^+$ are m/z=27 and m/z=18 amu (see Tab.C.1 in appendix C). The first one is associated with the fragmentation channel shown in the left part of Fig.9.8: $COOH^+ + NH_3CHCH_2^+ \rightarrow$

$COOH^+ + NH_3 + CHCH_2^+$. This channel appears in competition with the one related to $(18^+/45^+)$ i.e. $COOH^+ + NH_3CHCH_2^+ \rightarrow COOH^+ + NH_4^+ + CHCH$. The latter explains the correlation $18^+/45^+$ (barrier on the PES at 20.74 eV). Both coincidence measurements $(18^+/45^+$ and $27^+/45^+)$ show very similar relative intensities (see Tab.C.1 in appendix C). Accordingly, both channels require similar fragmentation energies (Fig.9.8). The pairs $(28^+/45^+)$ and $(27^+/45^+)$ correspond to the loss of two or three hydrogen atoms from $NH_2CH_2^+$. The theoretical slopes of the three pairs $(27^+/45^+)$, $(28^+/45^+)$ and $(30^+/45^+)$ are respectively -1.63, -1.57 and -1.47, which is also in agreement with the experimentally measured slopes (see Tab.C.1 in appendix C). They correspond to Coulomb explosion with $C_{carboxylic} - C_\alpha$ bond breaking followed by emission of neutral moieties. Further fragmentation of $NH_3CHCH_2^+$, proposed in Fig.9.8, provides the mechanisms to explain the peaks related to m/z=15, 16 and 17 amu in coincidence with $COOH^+$(m/z=45 amu). In the same figure we can also observe the pathway leading to the quite stable protonated ethylenimine (EI) cyclic cation ($NH_2CH_2CH_2^+$), which is a three-membered ring derivative of the monomer employed in the polymerisation by a cationic stepwise mechanism of the imino group.[258] Moreover, it is well known that ethylenimine and its derivatives cause degradation of DNA[259] and in reaction with 2-deoxy-D-ribose produce H_2O_2, a strong oxidative compound.[258] It is thus reasonable to propose the protonated EI cation as a dangerous and active compound produced with a relatively strong intensity in the fragmentation of β-alanine dication.

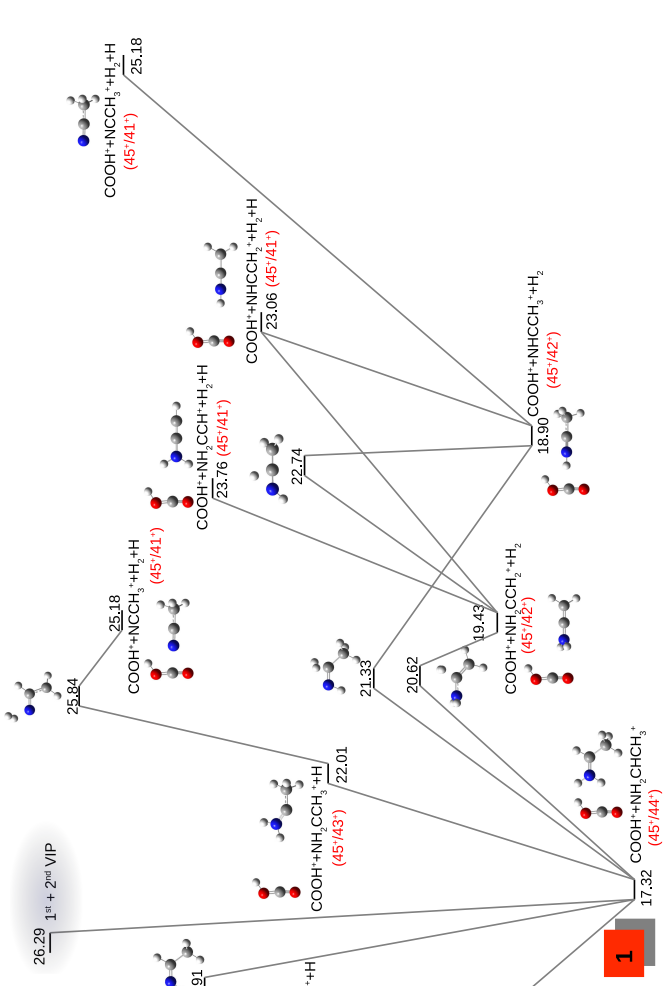


Figure 9.7: Fragmentation pathways for doubly positively charged β-alanine - Stationary points of the potential energy surface corresponding to $C_{carboxyl} - C_{\alpha}$ bond breaking and further fragmentation pathways leading to charged fragments: m/z=45 amu in coincidence with other charged fragments (m/z=43, 42, 41 amu). Relative energies, ΔE, are given in eV at the DFT-B3LYP/6-311++G(d,p) level of theory including the zero point energy correction with respect to the most stable isomer of the neutral molecule. Correlation notation is also included, *e.g.* the pair ($44^+/45^+$).

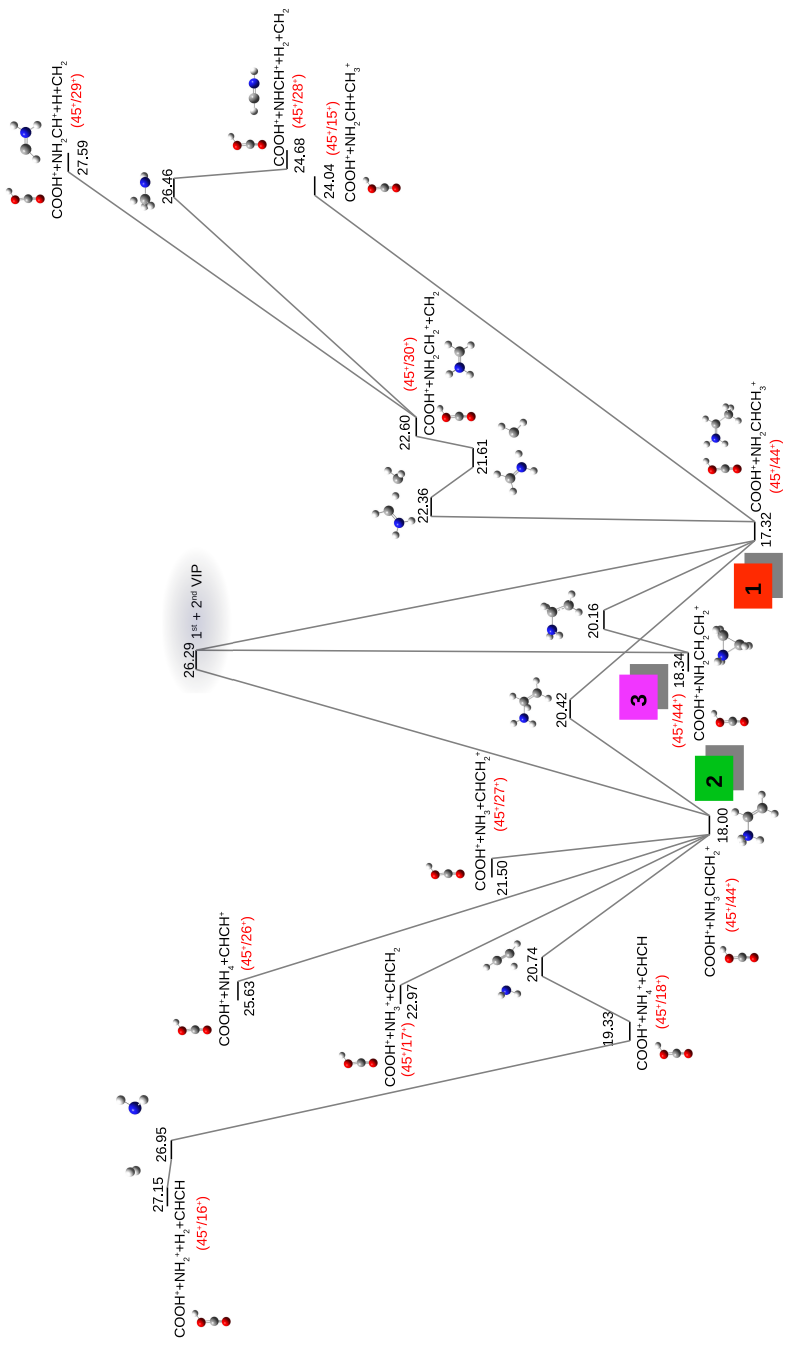


Figure 9.8: Further fragmentation due to $C_{carboxyl} - C_{\alpha}$ bond breaking are shown leading to charged fragments: m/z=45 amu with other heavy fragments (m/z=30, 28, 27, 26, 18, 17, 16 and 15 amu). Detailed information in the caption of Fig.9.7.

(b) $C_\alpha - C_\beta$ bond breaking.

$C_\alpha - C_\beta$ bond cleavage gives $NH_2CH_2^+ + CH_2COOH^+$ (expected coincidence $30^+/59^+$). However, in the experiment we do not observe the signal at m/z=59 amu (see Fig.9.1). This cation can easily break by losing CH_2 or OH without barriers. Indeed, the relatively low energies calculated for the fragments produced in the corresponding pathways, $HOCH_2CO^+$ or CH_2COOH^+, and the absence of a peak at m/z=59 amu in the experiments, allow us to conclude that the fragment at m/z=59 amu undergoes further fragmentation, in these kinetically favourable processes, leading to signals at $(30^+/42^+)$ and $(30^+/45^+)$ (see Fig.9.9 for the corresponding mechanisms). The observation of the signal at $(30^+/42^+)$ also indicates that the molecule has enough internal energy to break leading to $NH_2CH_2^+ + CH_2CO^+ + OH$. These species can be obtained in two different pathways. (i) OH emission followed by Coulomb repulsion [β-alanine$^{2+} \rightarrow NH_2CH_2CH_2CO^{2+} + OH \rightarrow NH_2CH_2^+ + CH_2CO^+ + OH$], (ii) first Coulomb explosion [β-alanine $^{2+} \rightarrow NH_2CH_2^+ + CH_2COOH^+$] and then emission of neutral OH [$NH_2CH_2^+ + CH_2CO^+ + OH$]. The associated computed island slopes are -1 and -0.71, respectively. The experimentally measured slope suggests first emission of neutral OH leading to $NH_2CH_2CH_2CO^{2+}$ followed by Coulomb repulsion producing $NH_2CH_2^+ + CH_2CO^+$ (see pathway in Fig.9.9).
The right part of Fig.9.9 shows that the evaporation of CH_2 from CH_2COOH^+ leads to the coincidence $(30^+/45^+)$. Further fragmentation of $NH_2CH_2^+$ by emission of H_2 leads to $NHCH^+$ in coincidence with $COOH^+$, island $(28^+/45^+)$. The most intense island observed in the experiment is m/z=30 amu ($NH_2CH_2^+$) in correlation with m/z=42 amu (CH_2CO^+) coming from $C_\alpha - C_\beta$ bond breaking. For the first fragment we observe similar many body fragmentation processes as the one described above, *i.e.* loss of one or two hydrogen atoms from $NH_2CH_2^+$ leading to the pairs $(29^+/42^+)$ and $(28^+/42^+)$, respectively. In Fig.9.6b, we can see that the three pairs have an experimental slope very close to -1. This is an indication that the molecular dication β-alanine^{2+} first loses a neutral OH fragment and then splits into two singly charged cations. Another group of islands involving the most prominent cation $NH_2CH_2^+$ is observed in correlation with the $COOH^+$ fragment.

Interesting molecular structures are observed for the spectra associated with the peaks at m/z=44, 43, 42 and 41 amu (see Fig.9.10). These cations present resonant structures with π delocalised electrons or strong N = C, N $\equiv$ C bonds. The associated exit channels appear a few eV below the entrance channel due to the high stability of such species (see Fig.9.7). However, this trend is not observed for the row at m/z=30, 29 and 28 amu (see Fig.9.6b). The lack of the peak at m/z=29$^+$amu in the experiment is caused by the fact that this molecule is a radical and its relative stability is much lower than the corresponding closed shell molecules (m/z=30, 28 amu) (see Fig.9.9).

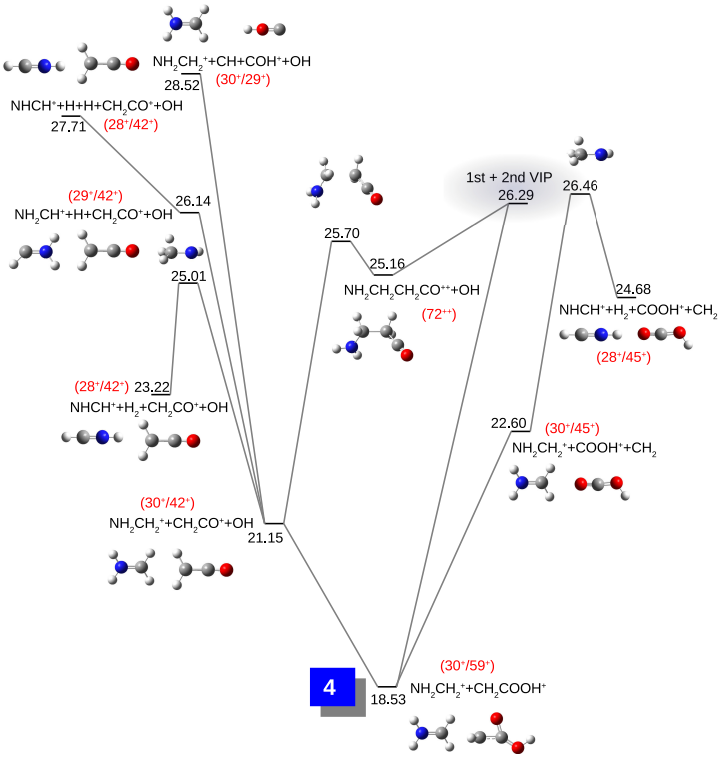


Figure 9.9: Stationary points of the potential energy surface corresponding to $C_\alpha - C_\beta$ bond breaking. Correlation notation is included, *e.g.* $(44^+/45^+)$. Detailed information in the caption of Fig.9.7.

(c) H migration.

We now focus on the exit channels produced after the most important isomerisation mechanism observed in the dynamics simulations: formation of the doubly charged geminal diol $NH_2CHCH_2C(OH)_2^{2+}$. In particular, we analyse two mechanisms: (i) Coulomb explosion after H transfer (see Fig.9.11) and (ii) emission of neutral fragments leading to stable dicationic species (see Fig.9.12). Both mechanisms are in competition with the rest of the processes but occur with lower probability, similar to our previous observations for glycine.[176]

Fragmentation of diol^{2+} due to the Coulomb explosion implies three exit channels: $NH_2CHCH_2CO^+ + H_2O^+$ $(71^+/18^+)$, $NH_2CHCHCOH^+ + H_2O^+$ $(71^+/18^+)$ and $NH_2CHCHCO^+ + H_3O^+$ $(70^+/19^+)$. A subsequent fragmentation pathway of $NH_2CHCH_2CO^+$ has also been calculated leading to $NH_2CHCH_2^+$ $(43^+/18^+)$ with the emission of a neutral carbon monoxide molecule. Fig.9.11 shows the corresponding pathways in the PES. All of them are below the entrance channel but

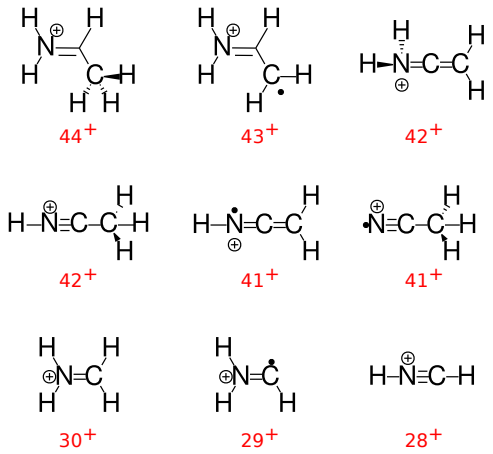


Figure 9.10: Schematic structures of molecular cations with m/z = 44, 43, 42, 41, 30, 29 and 28.

they can be produced only if one considers atomic rearrangement (even double H transfer) before fragmentation. The corresponding correlation islands present low relative intensity in the experiment and, accordingly, in the AIMD simulations the probability to occur is very small.

Hydrogen migration can be also followed by emission of neutral fragments. Indeed, the stable small dicationic molecules detected in the experiment are produced after emission of neutral species: H_2O, H, H_2, CH_2, CO, CO_2 or CH_2CO from doubly charged diol. The study of this part of the PES is presented in Fig.9.12 and it explains the different chemical mechanisms leading to molecular dications experimentally detected at signals: m/z= 35.5, 35, 34.5, 34, 21.5, 21, 20.5, 20 and 14.5 amu (see zooms-in in Fig.9.1). The pathways explaining these peaks come from the key structure $NH_2CHCH_2CO^{2+}$ produced after neutral H_2O loss from the doubly charged diol, through the formation of the intermediate, weakly bonded complex $NH_2CHCH_2CO\cdots H_2O^{2+}$. $NH_2CHCH_2CO^{2+}$ is a very stable acylium ion showing a resonant structure. The dehydration mechanism does not involve the amine group[49] and energetically is more preferable than splitting the charge into two fragments. This is the reason why in the experiment we almost do not observe the pair $(71^+/18^+)$ in the coincidence map (even if this channel appears at lower energy than the entrance one). Starting from the key structure $NH_2CHCH_2CO^{2+}$ the loss of H, H_2, CO and CH_2CO explain the peaks at 35, 34.5, 21.5 and 14.5 amu: $NH_2CHCHCO^{2+}$, $NHCCH_2CO^{2+}$, $NH_2CHCH_2^{2+}$ and NH_2CH^{2+} respectively. A subsequent emission of H and H_2 from these dications explain the rest of peaks corresponding to doubly charged molecules: (m/z=34, 21, 20.5 and 20 amu). Moreover, charge separation of the key structure $NH_2CHCH_2CO^{2+}$ can explain the island at $29^+/42^+$ (two body charge separation slope equals -1): (β-

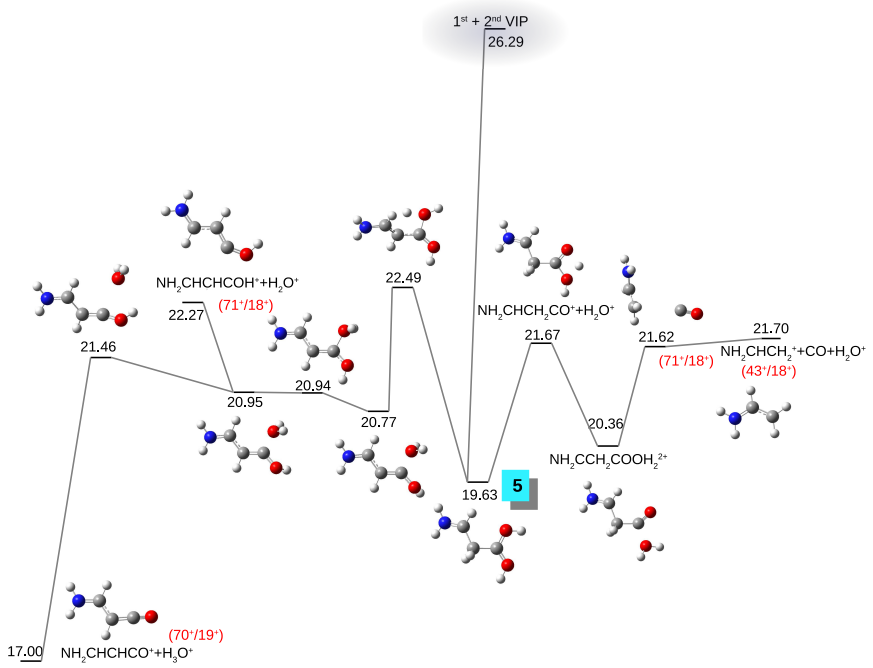


Figure 9.11: Stationary points of the potential energy surface corresponding to H migration and further Coulomb explosion, leading to two charged fragments. Correlation notation is also included, *e.g.* $(71^+/18^+)$. Detailed information in the caption of Fig.9.7.

alanine)$^{2+} \rightarrow$ diol$^{2+} \rightarrow NH_2CHCH_2CO^{2+} + H_2O \rightarrow NH_2CHCH^+ + COH^+ + H_2O$. The study of the PES also predicts an alternative mechanism for the formation of $NH_2CHCH_2^{2+}$ (m/z=21.5 amu), avoiding dehydration from the weakly bounded intermediate complex. The dication of diol^{2+} also undergoes dehydrogenation and simultaneously decarboxylation in a one-step three body fragmentation mechanism: $NH_2CHCH_2C(OH)_2^{2+} \rightarrow NH_2CHCH_2^{2+} + CO_2 + H_2$. The exit channel corresponding to this mechanism is energetically more favourable than most of the previous ones proposed to produce dications. Further subsequent fragmentation of the produced doubly charged molecule ($NH_2CHCH_2^{2+}$) by loosing H, H_2, CH_2 are alternative paths to explain the peaks at 21, 20.5, 20 and 14.5 amu, respectively.

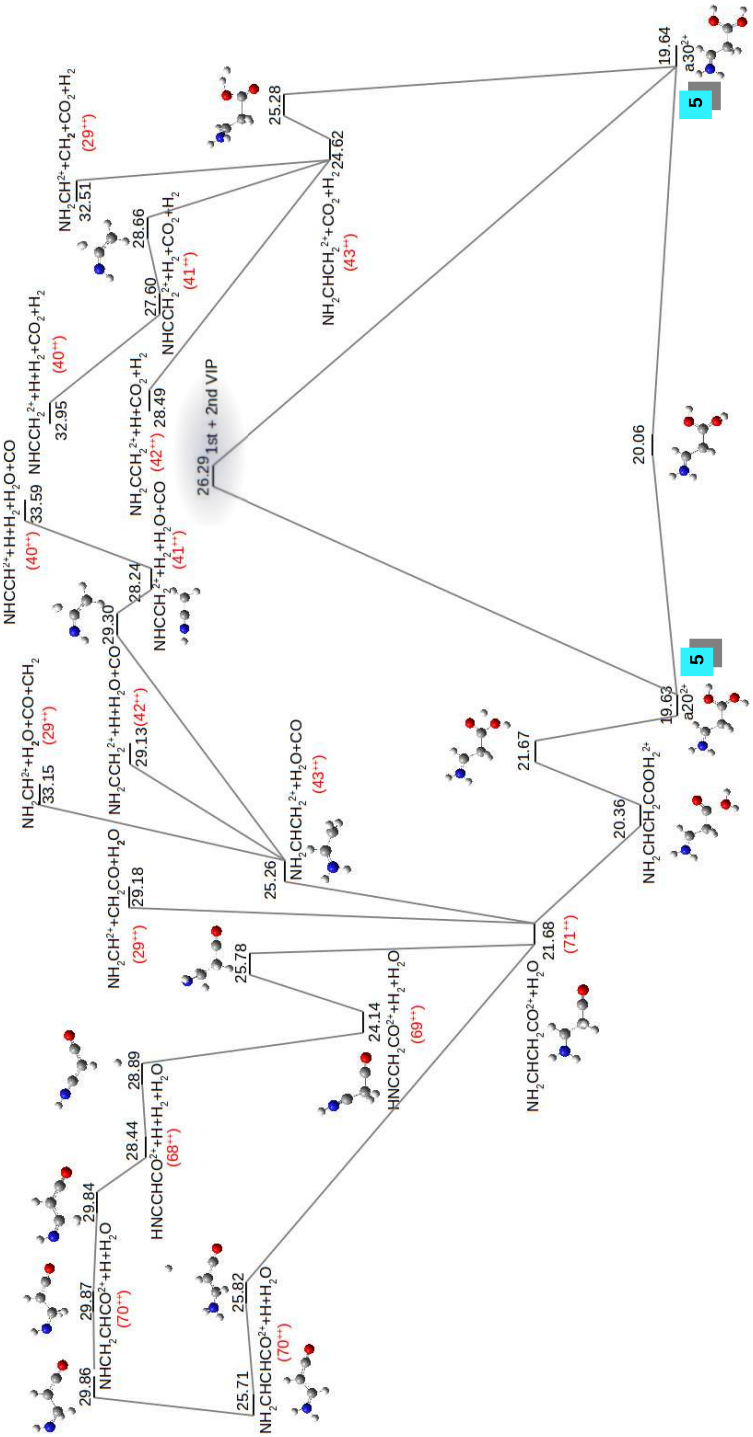


Figure 9.12: Stationary points of the potential energy surface corresponding to H migration and further emission of neutral species, leading to doubly charged fragments. Mass spectrum notation is also included, *e.g.* 71^{++} corresponds to m/z=35 amu. Detailed information in the caption of Fig.9.7.

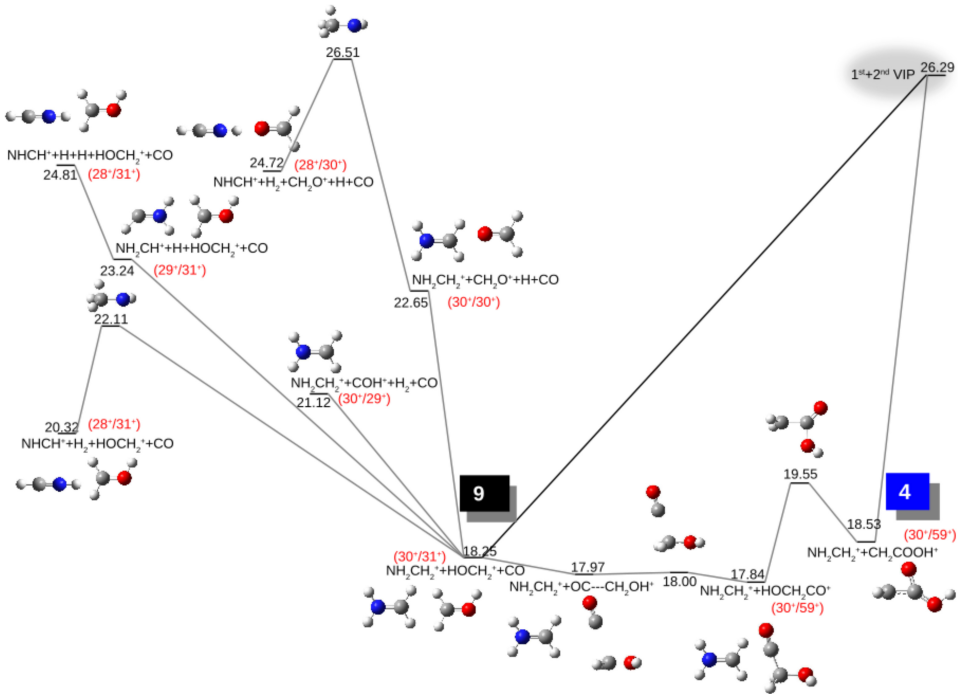


Figure 9.13: Stationary points of the potential energy surface corresponding to hydroxyl group migration and further fragmentation pathways. Coincidence notation is also included, *e.g.* $(30^+/59^+)$. Detailed information in the caption of Fig.9.7.

(d) OH migration.

In the ion-induced fragmentation of glycine, hydrogen migration was observed in competition with direct Coulomb explosion and was essential to understand the fragmentation dynamics.[176] For β-alanine our molecular dynamics simulations have shown a quite probable fragmentation pathway in competition with H migration and Coulomb explosion: hydroxyl group migration accompanied with emission of neutral CO and leading to two singly charged fragments ($NH_2CH_2^+$ and CH_2OH^+). This channel appears at very low energy in the PES (see Fig.9.13) and is assigned to the correlation $(30^+/31^+)$ measured with a very strong relative intensity (see Fig.9.6c and Tab.C.1 in appendix C). Moreover, from the entropic point of view it is more favourable than other channels. The computed mechanism reveals that the metastable $HOCH_2CO^+$ structure is involved before the loss of a neutral carbon monoxide and finally producing the detected ions: $NH_2CH_2^+$ and $HOCH_2^+$. The metastable structure observed in the molecular dynamics ($HOCH_2CO^+$) is an isomer of a CO protonated α-lactone.[260] These species were previously observed in the gas phase reactions of amino acids with the nitrosonium ion (NO^+).[261] Moreover,

previous studies have demonstrated that the open carbo-cation CH_2COOH^+ is not a saddle point on the PES.[262;263] AIMD simulations show that CH_2COOH^+ easily collapses to a CO protonated α-lactone (a transient species previously described in mass spectrometry experiments[264]), and the ring opens to form a more stable cation ($OCCH_2OH^+$). Further fragmentation of this cation leads to hydroxymethyl (CH_2OH^+) or to formaldehyde (CH_2O^+) cations. The latter one was also previously detected.[265] Interestingly, both species have been identified from spectral lines in the interstellar medium[266] and reactions between them can produce glycol aldehyde, acetic acid and methyl formate.[267] On the other hand, the hydroxymethyl cation (m/z=31 amu) was previously observed only in the single ionisation process after direct side chain bond breaking of serine.[169]

The fragmentation mechanism presented here is the only possibility to explain the coincidence island at ($30^+/31^+$). For the double ionisation process the driving force for this mechanism is the high stability of the products $NH_2CH_2^+ + HOCH_2^+ + CO$ (18.25 eV) with a relative energy 8 eV lower than the entrance channel and much lower than other channels with different charge distributions (*e.g.* $NH_2CH_2^+ + HOCH_2 + CO^+$ appears at 24.72 eV). The iminium cation ($NH_2CH_2^+$) possesses a strong hydride ion affinity.[268] Thus, in the first fs of the fragmentation the amino methyl cation can attract the hydride group from the other fragment CH_2COOH^+, causing the reorganisation of the geometry and allowing the movement of the OH group. This is indeed, the second most probable channel and leads to $NH_2CH_2^+ + CO + CH_2OH^+$ (see Fig.9.5b). Further fragmentation in this channel corresponds to the emission of neutral hydrogen molecule from $NH_2CH_2^+$ and leads to the coincidence measurement with $28^+/31^+$ ($NHCH^+/HOCH_2^+$). The presence of the pairs ($30^+/31^+$) and ($28^+/31^+$) (see Fig.9.6c.) can not be explained by the direct bond fragmentation of the molecular dication β-alanine^{2+} and must involve a rearrangement after Coulomb explosion. This OH migration is about three times slower than ultrafast H transfer and appears at around 120 fs of the simulations (see Fig.9.4). Moreover, this mechanism can be considered as one of the dominant at longer timescales due to the subsequent fragmentation of the products of mechanism (4). The exit channel after process (9) is energetically more favourable than the exit channel after process (4). Thus, at longer simulation time, products from (4) can easily evolve to products from (9) (see Fig.9.13). Indeed the second most intense coincidence island corresponds to the products of mechanism (9): $NH_2CH_2^+ + CH_2OH^+$ ($30^+/31^+$).

9.4 Conclusions

We have studied the fragmentation dynamics and different de-excitation pathways characterising the decay of excited doubly-charged β-alanine molecules in the gas phase. Experimentally, excited molecular dications were produced in collisions with O^{6+} ions and the fragments were identified by coincidence time-of-flight mass spectrometry allowing to determine the fragmentation dynamics and highlighting

the presence of new ionic species. Theoretically, we have found in *ab initio* molecular dynamics simulations that three types of processes occur in competition: (i) Coulomb explosion leads to the detection in correlation of two singly charged fragments; (ii) the formation of small stable dications produced by evaporation of neutral fragments following ultrafast intramolecular hydrogen migration (~ 40 fs); (iii) Coulomb explosion followed by hydroxyl group migration. The third process has not been previously reported and is expected to occur in other biomolecules. The good agreement between our simulation in the electronic ground state with the experimental measurements also sheds light on the charge state and excitation energy of the molecule after ionisation with highly charged ions and allows us to confirm that the excitation energy is quickly distributed over the vibrational degrees of freedom of the molecule.

Chapter 10

Results and Discussion

Stability and fragmentation of doubly- and triply-charged γ-aminobutyric acid

This chapter presents a combined experimental and theoretical study of the fragmentation of multiply-charged γ-aminobutyric acid molecules ($GABA^{z+}$, $z = 2, 3$) in the gas phase. The combination of ab initio *molecular dynamics simulations with multiple-coincidence mass spectrometry techniques allows us to observe and identify doubly-charged fragments in coincidence with another charged moiety. The present results indicate that double and triple electron capture lead to the formation of doubly-charged reactive nitrogen and oxygen species (RNS and ROS) with different probability due to the different charge localisation and fragmentation behaviour of $GABA^{2+}$ and $GABA^{3+}$. The MD simulations unravel the fast (femtosecond) formation of large doubly charged species, observed in the experimental microsecond timescale. The excess of positive charge is stabilised by the presence of cyclic $X-$member ($X = 3-5$) ring structures. 5-member cyclic molecules can sequentially evaporate neutral moieties, such as H_2, H_2O and CO_2 leading to smaller doubly charged fragments as those observed in the experiments.*

The results presented in this chapter has been published:
D. G. Piekarski, R. Delaunay, A. Mika, S. Maclot, L. Adoui, F. Martín, M. Alcamí, B. A. Huber, P. Rousseau, S. Díaz-Tendero and A. Domaracka, *Physical Chemistry Chemical Physics*, **accepted** (2017), DOI:10.1039/C7CP00903H.

10.1 Introduction

Reactive species in biochemistry such as multiply-charged ions and radicals play an important role in many processes including biomedical applications. For instance Reactive Oxygen Species (ROS) and their closely related Reactive Nitrogen Species (RNS) are key actors in oxidation-reduction reactions in biology and are as-

sociated with ageing and different diseases as cancer or diabetes.[269] In cancer, ROS are known to damage DNA[270;271] but are also involved in the chemical mechanisms of different cancer therapies,[272;273] for instances they are generated in ionised gas plasmas to be used in therapeutic applications.[274] Multiply-charged species are also important in biology and medicine: in radiotherapy multiple ionisation is less probable than other events (for instance electronic excitations or electron attachments) but produces more lethal effects to cells.[13;14] These multiply-charged molecular ions can also play an important role as precursors of new reactive species. In astrophysical environments multiply-charged ions are directly involved in the formation mechanism of unusual species,[275] and in biological media they can be formed and stabilised as the result of isomerisation-fragmentation reactions leading to species with a stable molecular structure and rather long lifetimes.[276]

The understanding of radiation damage at the molecular level requires a multiscale and multi-disciplinary approach.[277] Indeed the radiation damage starts within the first few fs during the passage of the impinging particle, at the so-called physical stage. In the case of collisions with "heavy" charged particles, such as ions (e.g. protons, carbon),[278] ionisation and direct damage (knock-out of atoms) can occur at this stage. At longer timescales, the resulting charged species and molecular fragments could lead to the release of radical species, which starts the so-called chemical stage.[279] As the body is mainly composed of water, its radiolysis and the associated formation of radical species (e.g. hydroxyl, hydroperoxyl and peroxide) are one of the main cause of secondary effects in radiation damage.[280] Thus, the study at the molecular level of the ionisation and fragmentation of the molecular systems is a fundamental step for the understanding and improvement of cancer treatment.[206]

In Chaps. 8-9, combined experimental-theoretical studies we have shown the formation of dications by molecular dissociation of amino acids after the interaction with slow multiply-charged ions in the gas phase.[176;177] Such collisions were used to efficiently produce multiply charged amino acid ions that could suffer promptly, in few tens of fs, a molecular rearrangement in order to stabilise the molecule against the coulombic repulsion. The formed doubly-charged isomers can further dissociate through "unexpected" fragmentation channels such as the emission of neutral water from the diol form of glycine[176], or CO-loss after OH-migration in $\beta-$alanine.[177] Thus, the processes involving rearrangement followed by fragmentation can lead to the formation of dicationic species stable at the ms timescale.

In this chapter, we study the fragmentation dynamics of the $\gamma-$aminobutyric acid (GABA), using this coupled experiment-theory approach. GABA is the simplest γ amino acid and is the main inhibitory neurotransmitter.[281] Moreover, GABA also appears as a cancer-signalling molecule as it plays a significant role in cell proliferation.[282] GABA was studied before in our group in collisions with low-energy O^{6+} ions.[92] In that study no intact $GABA^{2+}$ was observed because it spontaneously dissociated due to Coulomb explosion. All fragmentation channels of doubly-charged GABA were explained by exploration of the Potential Energy Surface (PES) using Density Functional Theory (DFT) considering only direct C−C cleavages and

charge splitting in the produced fragments. In the present study, results of new collision experiments using Ar^{9+} ions are presented. The main new aspect is the formation of a large variety of stable doubly charged species that were not observed in the previous experiments. These doubly charged species can only be formed by mechanisms not considered before[92] as isomerisation processes in doubly charged GABA followed by neutral emissions. Although these channels are less important than the Coulomb explosions already studied,[92] they represent important pathways for the formation of new highly-reactive ROS and RNS. Moreover, in these new experiments, stable triply-charged species are also formed and explain some of the observed measurements. Therefore, the present work will focus on the formation of reactive species by rearrangement-dissociation of GABA dications and trications. Experimental observations will be rationalised by Quantum Chemistry calculations based not only on explorations of the PES, but also on Molecular Dynamics (MD) simulations. MD simulations are necessary to describe the fast evolution of multiply-charged species and to locate the transient isomers formed from multiply positively charged GABA molecules.

10.2 Methods

Computational Details

Quantum chemistry calculations were carried out using the density functional theory (DFT). In particular, *ab initio* molecular dynamics calculations (AIMD) were performed using the Atom-centered Density Matrix Propagation method (ADMP),[141;142;149;150] employing the B3LYP (Becke's three parameter non-local hybrid exchange potential[114] with the non-local correlation[215]) functional together with the 6-31++G(d,p) basis set. The simulations were carried out using a time step of $\Delta t = 0.1$ fs, a fictitious mass of $\mu = 0.1$ amu and a maximum simulation time of $t_{max} = 450$ fs. To mimic the ion-molecule interaction, two or three electrons were extracted from the highest occupied molecular orbitals (HOMO) and (HOMO+HOMO-1) in a Franck-Condon type transition, and a certain amount of excitation energy ($E_{exc} \approx 0.03 - 4.08$ eV) was given at the same time. This internal energy was randomly distributed over all the vibrational degrees of freedom in each trajectory and corresponds to internal vibrational temperatures of $T_{vib} \approx 20 - 2300$K. In this way, we reproduced the sudden ionisation and excitation process produced in the gas phase collision of ions with neutral GABA molecules. 20 trajectories of AIMD simulations were carried out for 5 values of the excitation energy. The 30 most stable conformers of neutral GABA in the gas phase were considered,[283] running finally 3000 trajectories for the dication and trication. Statistics were then performed over all these trajectories. The potential energy surface (PES) was explored at DFT-B3LYP level of theory with 6-311++G(d,p) basis set. This method has been shown to describe accurately the PES of cationic amino acids.[72;92;164;169;176;177;238] All the calculations were performed using the Gaussian09 program.[195]

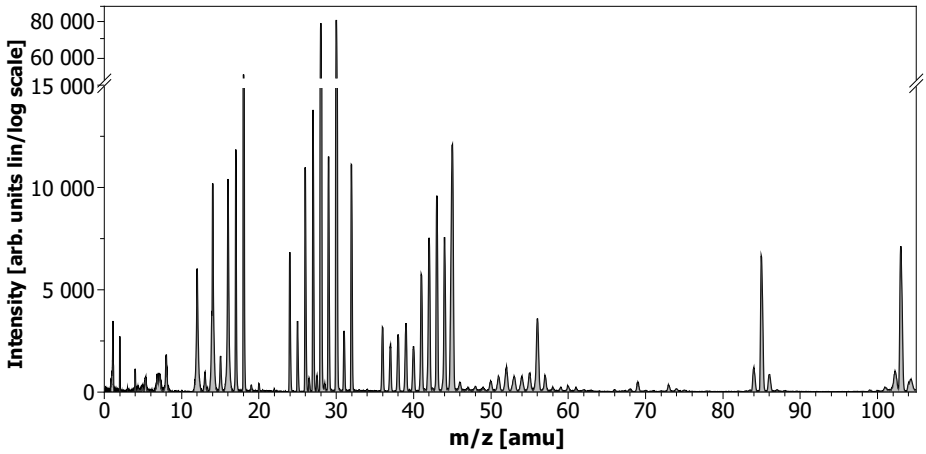


Figure 10.1: Total mass spectrum of the cationic products of the interaction of 135 keV Ar^{9+} ions with GABA molecules.

Experimental details

The interaction of 135 keV Ar^{9+} ions with GABA molecules in the gas phase was studied by analysis of the cationic products using coincidence time-of-flight mass spectrometry. The experiments were performed at ARIBE, the low-energy ion beam facility of GANIL in Caen, France.[214] The crossed beams experimental setup has been described in detail elsewhere,[188] thus only a brief description is given here. The Ar^{9+} ion beam was extracted from the electron cyclotron resonance ion source, mass selected, pulsed, collimated and transported into the interaction zone, where the beam crossed a molecular effusive jet of neutral GABA (Sigma-Aldrich, purity $\geq$ 99%) molecules evaporated from a heated oven device (T=375 K). Cationic products of the interaction were extracted into a Wiley-McLaren linear time-of-flight mass spectrometer.[213] Their arrival times were digitised in an event-by-event mode producing a list of events. They can be classified with the number (n) of detected charged fragments (n STOP). Thus, the charge state of the molecule before fragmentation is deduced and correlations between the emitted fragments are obtained (for more details see[92]).

With the ion-ion coincidence measurements we obtain detailed information about the fragmentation processes at short timescales (microsecond). An alternative experimental approach that provides complementary information at longer timescales (millisecond) is the tandem mass spectrometry (MS/MS). In relation with the molecule of interest in this chapter, electrospray ionisation tandem mass spectrometry (ESI-MS/MS) was employed to study the fragmentation patters of the GABA derivative N-diisopropyloxyphosphoryl-γ-amino butyric acid (DIPP-γ-aba)[284] and the reactivity of DIPP-γ-aba with nucleosides.[285] The same tech-

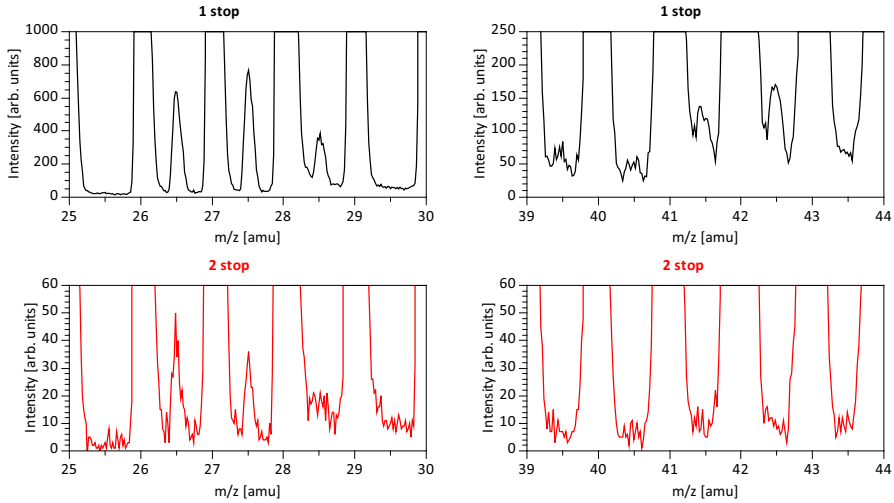


Figure 10.2: Zoom-ins of the 1 STOP (upper panel) and 2 STOP (lower panel) mass spectrum present the cationic products after the interaction of Ar^{9+} ions with neutral GABA at the energy of 135 keV underlining the presence of doubly charged fragments at m/z = 26.5, 27.5, 28.5, 39.5, 40.5, 41.5 and 42.5 amu.

nique (ESI-MS/MS) was also used to understand other chemical processes such as carbonyl oxygen migration in the sodium adduct of N-diisopropyloxyphosphoryl (DIPP) dipeptide methyl ester of GABA (DIPP-γ-amino butyric acid-Phe-OMe)[286] or to elucidate the mechanism behind the synthesis of GABA analogues.[287]

10.3 Results and discussion

The total mass spectrum of positively charged products from collisions of Ar^{9+} at 135 keV ions with neutral GABA molecules in the gas phase is presented in Fig.10.1. Similarly to previous investigations with 48 keV O^{6+} collisions,[92] the mass spectrum is dominated by molecular fragmentation. However, in the present studies the signal corresponding to the molecular parent cation is thrice stronger. Atomic collisions are characterised by a wide distribution of projectile impact parameters thus leading to an energy deposit distribution.[97;180;193;198] In the case of keV multiply charged ion collisions, the dominant process is electron capture by the projectile where at least one electron of the molecule is transferred to the incoming projectile. The capture radius depends on the projectile charge state and higher charge states are associated with a larger radius. Thus, the interaction is more gentle and the excitation energy transferred to the system is reduced. Therefore, in the case of highly charged projectiles, colder cations are produced which can survive the collision with higher probability.[251]

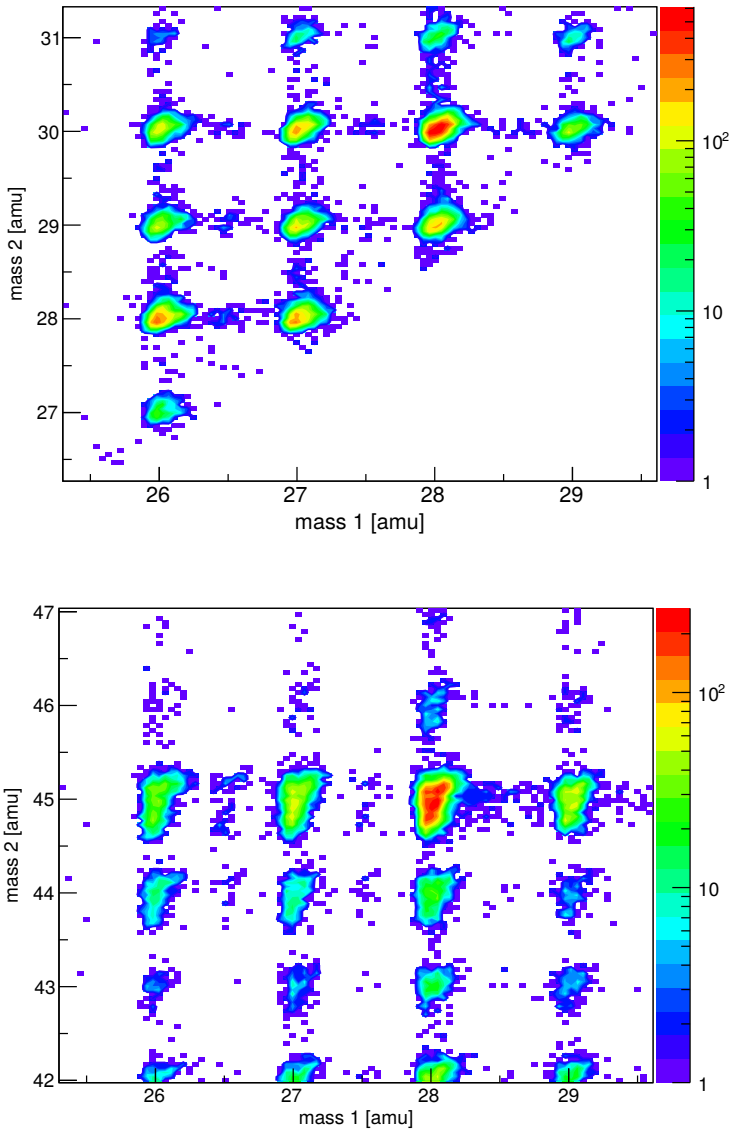


Figure 10.3: Coincidence map reflecting the fragmentation of the dication and trication of GABA molecules into two charged fragments after the double and triple electron capture with 135 keV Ar^{9+} ions, respectively. The coincidence of the interesting doubly charged fragments are shown: m/z = 26.5, 27.5, 28.5 amu (mass 1) is plotted as a function of the singly charged one; upper panel: 28, 29 and 30 amu (mass 2); lower panel 44 and 45 amu (mass 2).

In order to determine the initial charge state of the fragmented GABA molecule, we analyse the mass spectra associated with n STOP events ($1 \leq n \leq 3$). The 1 STOP mass spectrum is created from events where only one charged fragment has been detected after the collision. Thus, in view of the high detection efficiency, it mainly reflects the fragmentation of singly charged $GABA^{+}$ molecules.

Surprisingly, we observe strong signals at m/z = 26.5, 27.5, 28.5, 39.5, 40.5, 41.5 and 42.5 amu in the 1 STOP mass spectrum (see Fig.10.2), which we assigned to doubly charged fragments $m^{z+} = 53^{2+}$, 55^{2+}, 57^{2+}, 79^{2+}, 81^{2+}, 83^{2+} and 85^{2+}, respectively. The presence of these fragments in the 1 STOP mass spectrum indicates that $GABA^{2+}$ may decay via the emission of neutral fragments. These species were not detected in previous experiments using O^{6+} projectiles[92]. Former studies on smaller amino acids as glycine[176] or $\beta-$alanine[177] detected doubly-charged species (peaks at 14.5, 27.5 and 28.5 amu in the case of glycine and 14.5, 20.5, 21.5, 34.5 and 35.5 amu in the case of $\beta-$alanine) and were explained by isomerisation mechanisms that stabilised them. It is important to note that in GABA larger doubly-charged fragments are observed. This is relevant because the larger the doubly-charged fragment the easier its stabilisation and the longer its lifetime, so that they can play a more important role as a secondary reactive species in radiation damage.

The 2 STOP mass spectrum is associated with events where two charged fragments were detected after the collision and it contains mainly information about $GABA^{2+}$ fragmentation into two singly-charged fragments. These are usually associated with Coulomb explosion.[176;177;211] However, we observe also doubly-charged fragments in the 2 STOP mass spectrum (see lower panels in Fig.10.2). These dications arise from the dissociation of $GABA^{z+}$ with $z \geq 3$. In the case of smaller amino acids, like glycine[176] or $\beta-$alanine[177], doubly-charged species were not observed in the 2 STOP mass spectrum or in the coincidence maps. From the 2 STOP mass spectrum we can plot a coincidence map which shows the correlation between the two detected charged fragments.[288] The obtained map is shown in Fig.10.3 and the quantitative values for the most relevant peaks are summarised in Tab.10.1. It is dominated by several intense islands corresponding to the correlation between two singly-charged fragments coming from the fragmentation of $GABA^{2+}$. The coincidence map also shows low-intensity islands corresponding to doubly-charged fragments measured in coincidence with singly-charged ones; mostly one cation with m/z = 30 or 45 amu and one dication with m/z = 26.5, 27.5 or 28.5 amu.

In order to clarify the different possible decay processes of doubly- and triply-charged GABA molecules, AIMD simulations were performed; for each charge state 3000 trajectories were considered. In section 10.3 we will present the fragmentation mechanisms after double electron capture and in section 10.3 the corresponding ones after triple electron capture. Finally in section 10.3, we will discuss the general mechanisms for the production of the doubly charged RNS and ROS species.

Table 10.1: Assignment of the correlation islands corresponding to $GABA^{2+}$ fragmentation associated to the computed exit channels. Relative experimental intensities (R. I.) in % are directly compared with the calculated relative energies (ΔE) in eV, given with respect to the $NH_2CH_2^+ + CH_2CH_2 + COOH^+$ channel, which is exothermic by 5.28 eV with respect to $GABA^{2+}$. The most intense correlation islands for (i) Coulomb explosion, (ii) Isomerisaton and (iii) Neutral emission are shown.

	m/z	R.I.	Assignment	ΔE (eV)
(i)	$30^+/45^+$	12.5	$NH_2CH_2^+ + CH_2CH_2 + COOH^+$	0.00
	$28^+/45^+$	6.3	$NH_2CH_2 + CH_2CH_2^+ + COOH^+$	3.85
			$NHCH^+ + H_2 + CH_2CH_2 + COOH^+$	2.03
	$28^+/30^+$	4.2	$NH_2CH_2^+ + CH_2CH_2^+ + COOH$	2.03
	$45^+/56^+$	1.2	$NHCHCH_2CH_2^+ + H_2 + COOH^+$	0.2
(ii)	$30^+/47^+$	0.06	$NH_2CH_2^+ + CHCH + HC(OH)_2^+$	-0.1
	$28^+/47^+$	0.03	$NHCH^+ + H_2 + CHCH + HC(OH)_2^+$	1.93
(iii)	$30^+/56^+$	0.9	$NH_2CH_2^+ + CH_3CHCO^+ + OH$	1.54
	$31^+/42^+$	0.4	$NHCCH_3^+ + H_2 + CH_2OH^+ + CO$	0.21

Double electron capture

In the MD simulations of the doubly-charged GABA molecule, 37 different fragmentation pathways have been characterised. Tab.10.1 and Figs. in the appendix C provide detailed information of the 37 pathways, statistics (number of the trajectories leading to a particular pathway) and the structure of the species more frequently obtained. They can be classified into three groups according to the key step in the trajectory: (i) Direct Coulomb explosions (CE), i.e. trajectories with a direct bond breaking from $GABA^{2+}$ leading to the separation into two charged fragments; (ii) Isomerisations: trajectories with an intramolecular H-transfer and the formation of $GABA^{2+}$ isomers; and (iii) Neutral emissions: trajectories that imply in their first steps the emission of a neutral fragment from $GABA^{2+}$ or one of its isomers. Figure 10.4a presents the global statistics of the results of MD simulations and shows how the three different processes compete as a function of the excitation energy for $GABA^{2+}$. As expected, the most probable process (close to 90% trajectories in the MD) corresponds to direct CE and consistently explains the most intense peaks in the coincidence map (see Fig.10.3 and Tab.10.1). Isomerisation plays a significant role even at low excitation energies and explains the formation of some observed ROS as doubly-charged geminal diol $-C(OH)_2$. Finally, emission of neutral fragments from $GABA^{2+}$ becomes more relevant at higher excitation energies and explains the doubly charged species observed in the spectra, which can present quite unusual structures for ROS and RNS, not considered in previous works.

Tab.10.1 illustrates the most intense correlation islands observed in the experiment. They are related to CE ($A^+/B^+ = 30^+/45^+$, $28^+/45^+$, $28^+/30^+$, and $45^+/56^+$), isomerisation to diols ($A^+/B^+ = 30^+/47^+$ and $28^+/47^+$) and emission of neutral moieties ($A^+/B^+ = 30^+/56^+$ and $31^+/42^+$). In general, for the CE

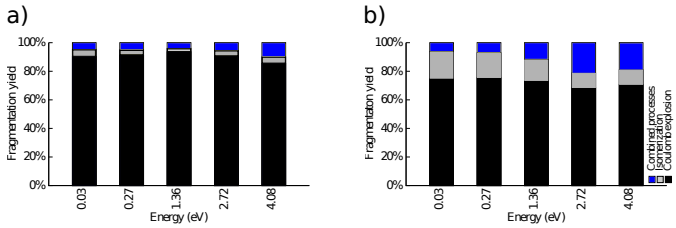


Figure 10.4: *Ab initio* molecular dynamics statistic for all 30 dications and trications of GABA. Competition between Coulomb explosion (i), isomerisation (ii) and combined processes (iii) is presented as a function of the excitation energy for a) $GABA^{2+}$ and b) $GABA^{3+}$.

processes the relative intensities are in good agreement with calculated relative energies (the higher the relative intensities the lower the relative energies). However, for more complex processes, as isomerisation or neutral emission, the PES values are not enough to explain the observed intensities. MD simulations give the complementary information about kinetically favourable processes. Detailed analysis of these processes can be found in the appendices and are summarised in the next subsections.

Direct Coulomb Explosions

CE in biomolecules is well known and has been previously observed for amino acids.[92;176;177;211] In previous experiments of collisions of GABA with O^{6+}, CE were the only channels found in the fragmentation of $GABA^{2+}$[92] and the observed fragments were inferred on the basis of the topology of the PES and the energy of the final products. Three different initial steps can be considered related to different cleavages of the carbon chain: $C_{carboxylic} - C_\alpha$, $C_\alpha - C_\beta$ and $C_\beta - C_\gamma$.

$C_\alpha - C_\beta$ cleavage represents the less favourable channel leading in the first step to $NH_2CH_2CH_3^+ + CH_2COOH^+$ (coincidence at $44^+/59^+$). These are quite unstable species that can further evolve to give $NH_2CH^+ + COOH^+ + CH_3 + CH_2$ ($29^+/45^+$) as described before.[92] The present MD simulations confirm this description: less than 10% of the trajectories follow this reaction path (see the appendices) and indeed the coincidence $44^+/59^+$ is not observed in the map and the $29^+/45^+$ coincidence is fairly weak.

Breaking the $C_{carboxylic}-C_\alpha$ bond leads to $NH_2CH_2CH_2CH_2^+ + COOH^+$ ($58^+/45^+$). For the former fragment the MD simulations show that at least 6 different isomers can be formed (see appendices). The most stable among them corresponds to $NH_2CHCH_2CH_3^+$, which is close in energy to a four-membered ring structure previously proposed.[92] These structures may lose H_2 giving the $56^+/45^+$ coincidence or split into $NH_2CH_2^+ + CH_2CH_2$ or $NH_2CH_2 + CH_2CH_2^+$ (coincidences $30^+/45^+$ or $28^+/45^+$ respectively, as the charge can be located in any of the two fragments).

The $30^+/45^+$ and $28^+/45^+$ coincidences may also be explained if the $C_\beta - C_\gamma$ bond breaks first leading to $NH_2CH_2^+ + CH_2CH_2COOH^+$ ($30^+/73^+$), and in a second step the cleavage of the $C_{carboxylic} - C_\alpha$ is produced. Most trajectories in the MD simulations follow this route. The first step is an exothermic process (5.08 eV). $CH_2CH_2COOH^+$ is a stable fragment that can adopt many isomeric forms (see the appendices), and is mainly detected in the 1 STOP measurements. However, the coincidence $30^+/73^+$ is not observed because the 73^+ fragment is produced with enough energy to further evolve giving e.g. $COOH^+$ (45^+ amu) or $CH_2CH_2^+$ (28^+ amu) fragments. This is consistent with the fact that the $30^+/45^+$ and $30^+/28^+$ coincidences are among the most prominent ones in the map. Still these two channels correspond to exothermic processes (by 5.28 and 3.17 eV respectively) and both $CH_2CH_2^+$ and $NH_2CH_2^+$ can further lose H or H_2 leading to the different islands with combination of masses between 30^+ and 26^+ observed in the coincidence map (upper panel of Fig.10.3). The energetics for these fragmentations are given in the appendices.

Among the fragments produced in the CE processes, we find quite reactive cationic or radical species such as $CH_2CH_2^+$, $NH_2CH_2^+$ or $COOH^\cdot$. Experimentally we get direct information by detecting the positively charged species and indirect information of neutral production by comparison with the simulations. Similar kinds of intermediates have been observed in photolysis,[289] by neutralisation-reionisation mass spectrometry experiments[290] and in fragmentation of molecules in excited electronic states.[291]

Isomerisations

Around 5% of the MD simulations correspond to isomerisation of $GABA^{2+}$. Final structures obtained in the trajectories were optimised and are given in the appendices. Most of the trajectories correspond to the formation of the geminal diol form $-C(OH)_2$. In Tab.10.1 two coincidence islands ($30^+/47^+$ and $28^+/47^+$ amu) are explained considering that isomerisation takes place before the fragmentation of the $GABA^{2+}$ molecule. Once the diol isomer is formed, the $C_{diol} - C_\alpha$ bond breaks leading to the above mentioned fragments. It has to be underlined that, most of the isomerisation leads to this tautomer of $GABA^{2+}$. Moreover, the small relative intensities measured for the islands $30^+/47^+$ (0.06%) and $28^+/47^+$ (0.03%) and associated with the production of $diol^{2+}$, are in agreement with the MD statistical analysis (see appendix D).

Neutral emissions

At low energies, 4% of the trajectories show the emission of a neutral fragment in the early stages of the dynamics and in many cases the formation of a stable doubly-charged fragment. At higher excitation energies, this percentage increases up to 10% and a wider variety of doubly-charged structures are obtained. These processes, schematically shown in Fig.10.5, play an essential role in $GABA^{2+}$ frag-

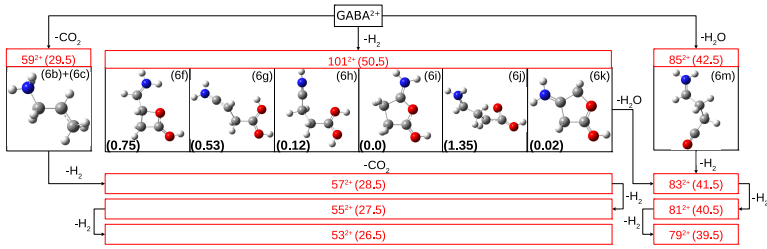


Figure 10.5: Cationic products of the fragmentation of doubly-charged GABA obtained in the MD simulations. Structures (6b+6c), (6f-k) and (6m) were optimised at DFT-B3LYP/6-311++G(d,p) level of theory. In parentheses the relative energies between conformers are given (in eV) with respect to the most stable isomer (6i). Possible structure of dications at m/z = 26.5, 27.5, 28.5, 39.5, 40.5, 41.5 and 42.5 amu are presented in the appendices.

mentation as they are necessary to explain the doubly-charged fragments observed in the 1 STOP spectrum in Fig.10.2.

According to the simulations, the first step corresponds to the loss of small neutral fragments such as CO_2, H_2, CO or H_2O. The most probable channels are the ones labelled as 6b and 6c (see Fig.10.5 and appendices) in which neutral CO_2 is emitted. The remaining doubly-charged molecule can be a stable fragment giving the observed 59^{2+} species (6b or 6c in Fig.10.5) or it can break into $NH_2CH_2^+ + CH_2CH_3^+$ (channel 6a) giving the $30^+/29^+$ signal in the coincidence map. The next most likely mechanism is the emission of H_2 giving several possible structures for the remaining doubly-charged fragment as diols or cyclic ethers forming 4- and 5-member ring structures (see Fig.10.5). Subsequent dehydrogenation, water loss or evaporation of CO_2 explains the experimentally observed doubly-charged fragments with masses 57, 55 and 53 amu. Dehydrogenations plus loss of water from diol isomers lead to the stable doubly-charged species with masses 83, 81 and 79 amu observed in the experiments. It is important to note that in these mechanisms all obtained structures with mass lower than 60 amu correspond to RNS species.

Triple electron capture

In the MD simulations of triply charged GABA molecules, we have identified 44 different fragmentation pathways which are classified into three groups, similarly to $GABA^{2+}$ (see Tab.D.2 in appendix D for the complete list): (i) Direct Coulomb explosions, (ii) Isomerisations and (iii) Neutral emissions and combined processes. This last category is slightly different from the $GABA^{2+}$ as it includes not only emissions of neutral fragments, and simple Coulomb explosions of the triply-charged species, but also complex Coulomb explosion that only occurs after a major isomerisation of GABA. It has to be taken into account that triply-charged species

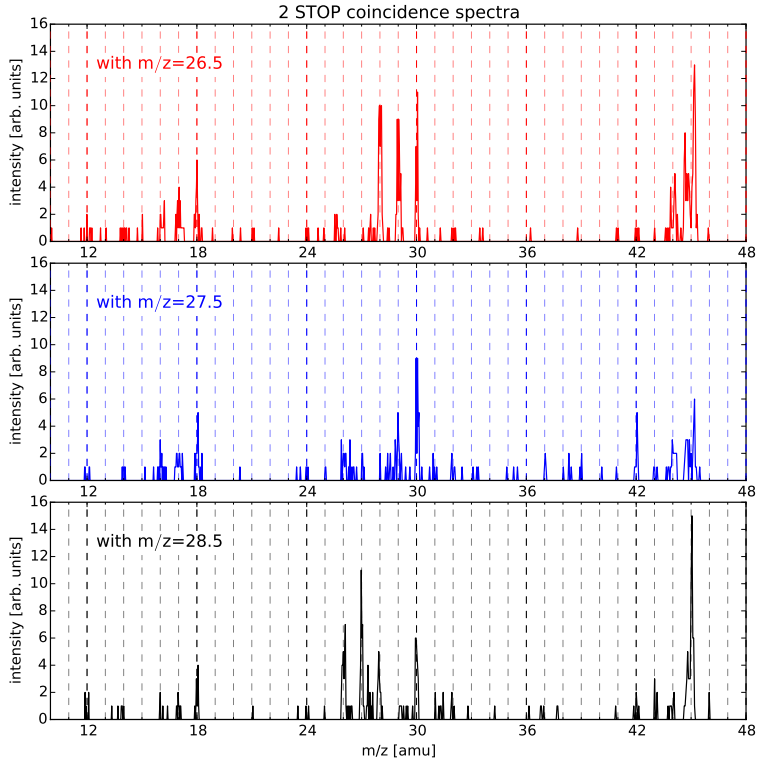


Figure 10.6: Mass spectrum for any cationic fragment in coincidence with doubly charged fragments at m/z = 26.5 (upper pannel), 27.5 (central pannel) and 28.5 (lower pannel) amu.

formed after neutral emissions are in general very unstable and evolve through bond cleavages and fragmentation into two charged fragments. The large Coulomb repulsion existing in triply-charged systems allows reaching complex fragmentation channels that we will identify as "combined processes". Figure 10.4b illustrates the probability of the different fragmentation mechanisms of $GABA^{3+}$ molecules as a function of the excitation energy. The exact percentage of each channel and the more relevant structures obtained in $GABA^{3+}$ fragmentation are also given in the appendices.

In the case of the $GABA^{3+}$ dissociation, besides the expected three body charge separation, we observe two body charge localisation producing a doubly-charged fragment in coincidence with a singly-charged moiety. Their correlations are given in Fig.10.6. Each of the doubly-charged fragments (m/z = 26.5, 27.5 and 28.5 amu) is observed in coincidence with a singly-charged one: 17^+, 18^+, 26^+, 27^+, 28^+, 29^+, 30^+, 42^+ and 45^+.

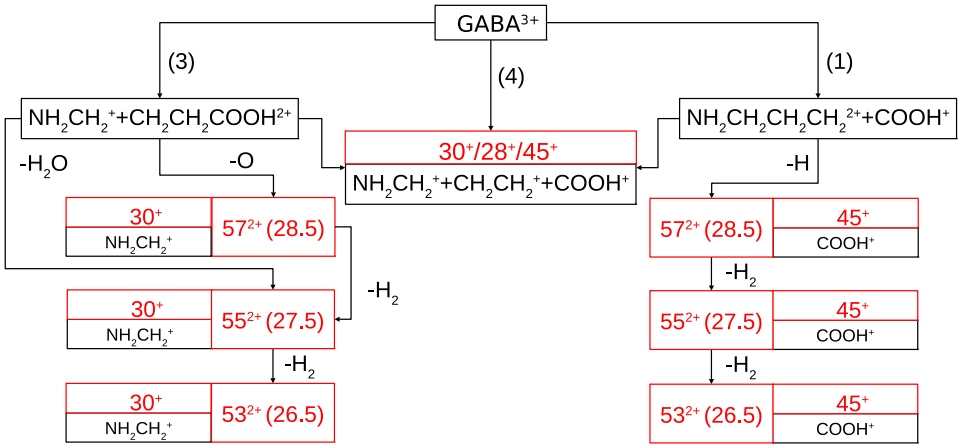


Figure 10.7: Scheme of $GABA^{3+}$ fragmentation based on the MD simulations (1), (3) and (4); For more details see the appendices. Further, possible emission of neutral fragments leads to dicationic species at m/z = 26.5, 27.5 and 28.5 amu measured in coincidence with $NH_2CH_2^+$ (m/z = 30 amu) and $COOH^+$ (m/z = 45 amu), left and right panel, respectively.

The fact that in the 3 STOP spectra coincidences with 30^+ and 45^+ are observed is very significant: these are common peaks observed in amino acid fragmentation, generally assigned to the iminium cation $NH_2CH_2^+$ and to the $COOH^+$ species, respectively,[92;176;177;210;211] which are consistent with the assignments done in this article for $GABA^{2+}$ fragmentation (see Tab.10.1). If the iminium cation is formed, the remaining doubly-charged species should be a ROS, while if the $COOH^+$ is emitted the doubly-charged species found in coincidence should correspond to a RNS. The fact that peaks 26.5, 27.5 and 28.5 are observed in coincidence with both 30^+ and 45^+ signals points out that these doubly-charged complexes are a mixture of RNS^{2+} and ROS^{2+} with the same masses. In the following subsection we will explore how these species can be formed.

It is important to realise that for the losses of neutral fragments discussed above for the fragmentation of $GABA^{2+}$, several mechanisms were observed in the MD simulations that lead to the same doubly-charged ions (m/z = 26.5. 27.5 and 28,5 see Fig.10.5) but in these cases the proposed mechanisms involved the formation of RNS^{2+} species only. In fact the dominant mechanisms imply loss of neutral CO_2. Therefore in the fragmentation of $GABA^{2+}$, formation of RNS^{2+} species is expected, while in the fragmentation of $GABA^{3+}$ both RNS^{2+} and ROS^{2+} should be formed. The structure of these doubly charged species can be quite complex as many isomeric forms are possible. They are summarised in appendix D, for RNS^{2+} and ROS^{2+}, respectively.

Direct Coulomb explosion

As shown in Fig.10.4b direct Coulomb explosions represent more than 75% of the fragmentation channels obtained in the MD simulations. As in the case of $GABA^{2+}$ three different bond breaking's can be envisaged: $C_{carboxylic}-C_\alpha$, $C_\alpha-C_\beta$ and $C_\beta - C_\gamma$. The fragmentation yield for each channel is given in Fig.7 of the appendices and a summary of the MD simulations showing the competition among the different channels is represented in Fig.10.7. Breaking of the $C_\beta - C_\gamma$ bond (channel 3 in Fig.10.7) represents the main channel, indeed 62.7% of the trajectories follow it, while $C_{carboxylic} - C_\alpha$ breaking (channel 1 in Fig.10.7) represents only 3% of the trajectories. In both cases further fragmentation may lead to the three-body Coulomb explosion ($NH_2CH_2^+ + CH_2CH_2^+ + COOH^+$), this is a fast process that is even observed in the time of the MD simulations for 5.2% of the trajectories.

The energy profile of these channels is given in Fig.10.8. Three charged moieties $30^+/28^+/45^+$ measured in coincidence correspond to the most intense correlation observed experimentally in $GABA^{3+}$ fragmentation. The contribution of the 3-body fission to the CE processes increases with the excitation energy (see the appendices). Also, Fig.10.8 shows that this CE process corresponds to the lowest-energy channel, 4 and 5 eV lower than the 2-body CE channels (1) and (3), respectively; the three processes are highly exothermic.

Alternative channels to the 3-body fissions can be also considered (see Fig.10.7). However, the fragment $CH_2CH_2COOH^{2+}$, formed in channel 1, or the fragment $NH_2CH_2CH_2CH_2^{2+}$, formed in channel 3, are not observed experimentally. These doubly charged fragments obtained in MD simulations may undergo further fragmentation. The islands $57^{2+}/45^+$, $55^{2+}/45^+$ and $53^{2+}/45^+$ observed in Fig.10.6 are indeed a consequence of the 2 body CE process (channel 1; formation of $NH_2CH_2CH_2CH_2^{2+} + COOH^+$) and subsequent H and H_2 losses (see Fig.10.7). In this case only RNS^{2+} species will be formed. The observation of $57^{2+}/30^+$, $55^{2+}/30^+$ and $53^{2+}/30^+$ coincidences in Fig.10.6 arise from the other 2 body CE process (channel 3: $NH_2CH_2^+ + CH_2CH_2COOH^{2+}$) followed by emission of light neutrals (O, H_2, or H_2O). In this path only ROS^{2+} species will be obtained. Figure 10.8 also shows the energy of the exit channel related to the smallest doubly-charged fragment m/z = 26.5 amu, which can correspond either to $NC_3H_3^{2+}$ or HC_3O^{2+} if channel 1 or channel 3 are followed. Both exit channels lie 10 eV below the entrance channel ($1^{st}+2^{nd}+3^{rd}$ vertical ionisation potential), i.e. the exothermicity of the first CE starting from $GABA^{3+}$ is enough to produce the fragments proposed in Fig.10.7.

The experimental relative intensities for the coincidence of doubly-charged fragments with the iminium ($NH_2CH_2^+$) and the carboxyl ($COOH^+$) cations are the same. One can thus conclude that formation of RNS^{2+} or ROS^{2+} are equally probable from $GABA^{3+}$. On the basis of this evidence, further analysis of the fragments corresponding to peaks at m/z = 26.5, 27.5 and 28.5 amu were performed (see appendix D for the details of their structure). It reveals that there exist many stable or metastable doubly-charged species, including cyclic structures such as 3-, 4- and

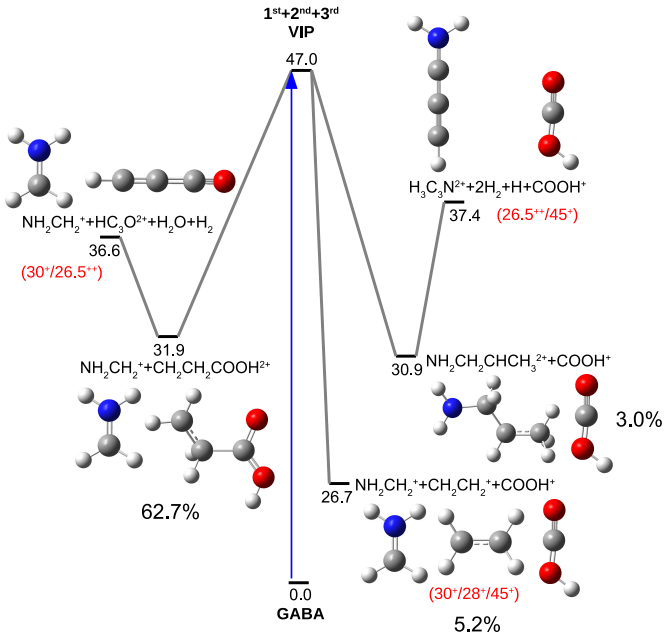


Figure 10.8: Fragmentation pathways for triply positively charged GABA. Only the stationary points of the potential energy surface corresponding to m/z = 26.5 amu (53^{2+}) in coincidence with $NH_2CH_2^+$ and with $COOH^+$ are presented. Relative energies, ΔE, are given in eV at the DFT-B3LYP/6-311++G(d,p) level of theory including the zero point energy correction with respect to the most stable neutral conformer of GABA. Numbers in % mean the fragmentation yield taken from MD simulations (see appendices for more details).

5-member ring structures, although the open chains are the most stable ones.

Isomerisation, neutral emissions and combined processes

MD simulations also show that a relevant number of trajectories in $GABA^{3+}$ (around 25%, see Fig.10.4), imply isomerisation or emission of neutral fragments. Most of the isomerisation processes involve a H transfer to the amino group NH_2 (∼10% of the total trajectories). From these structures, emission of NH_3^+ or NH_4^+ can be easily envisaged. In the coincidence spectra with doubly-charged fragments (see Fig.10.6), the additional contributions associated with peaks at m/z= 17 and 18 amu can be related to these isomerisations (which tipically take place in a few tens of fs[176;177])

$GABA^{3+} \xrightarrow{\text{H migration}} [NH_3 - R]^{3+}$

and subsequent emissions of NH_3^+ or NH_4^+, respectively. For the NH_3^+ case the possible scenario is:

$$[NH_3-R]^{3+} \xrightarrow[(17)]{-NH_3^+} CH_2CH_2CHCOOH^{2+} \xrightarrow{-CH_2CH_3} CCOOH^{2+}(28.5)$$

$$[NH_3-R]^{3+} \xrightarrow[(17)]{-NH_3^+} CHCH_2CH_2COOH^{2+} \xrightarrow[\text{OH migration}]{-CH_2OH} CHCH_2CO^{2+}(27.5)$$

$$[NH_3-R]^{3+} \xrightarrow[(17)]{-NH_3^+} CHCH_2CH_2COOH^{2+} \xrightarrow[\text{OH migration}]{-CH_2OH,-H_2} CHCCO^{2+}(26.5)$$

The analogue chemical reactions starting with the loss of NH_4^+ can explain the coincidence measurements $18^+/57^{2+}$, $18^+/55^{2+}$ and $18^+/53^{2+}$.

Other channels observed in the MD simulations involve neutral H_2 emissions as a first step and can also contribute to the observed coincidences in Fig.10.6. For instance one of the most probable channels (∼3% at low energies and ∼7% at high energies) implies H_2 emission, which followed by Coulomb explosion directly yields the $55^{2+}/45^+$ coincidence:

$GABA^{3+} \xrightarrow{-H_2} NH_2CCH_2CH_2COOH^{3+} \rightarrow COOH^+(45) + NC_3H_6^{2+}(27.5)$

Similar H_2 emissions, representing ∼1-2% of the trajectories lead to a different triply-charged intermediate, $NH_2CHCHCH_2COOH^{3+}$, which can evolve in the same way yielding the same coincidence $55^{2+}/45^+$. At high energies ∼3% of the observed trajectories directly correspond to H_2 emission and Coulomb explosion:

$GABA^{3+} \rightarrow NH_2CHCHCH_2^{2+} + COOH^+ + H_2$

Formation of triply-charged diol isomers is also obtained in the MD simulations, but always coupled with simultaneous emission of H_2, and they represent less than 1% of the observed trajectories. For instance the process:

$GABA^{3+} \rightarrow NH_2CCHCH_2C(OH)_2^{3+} + H_2$

is observed and could lead, by loosing water and CO, to a doubly-charged or even triply-charged RNS, but according to the statistics in the MD simulations these are very unlikely processes.

Formation of RNS^{2+} and ROS^{2+}

Multiply charged amino acids formed after interaction with multiply charged ion projectiles can evolve leading to reactive free radical species.[92;176;177;211;251] So formed oxidative chemical compounds can damage chemically important bio-compounds in the living cells e.g. they can work as starters of chain reactions resulting in oxidation of lipid termed as lipid peroxidation, protein oxidation and oxidation of nucleic acids like DNA or RNA.[292] These reactive species at higher concentration are well known mediators, important for damaging cell structures, including nucleic acid, proteins, lipids and membranes.[293]

The formation of highly reactive singly- and doubly-charged intermediates containing oxygen and nitrogen seems to be common in the fragmentation of GABA. The results of the MD simulations summarised in Figs. 10.5 and 10.7 and those of the experimental coincidence measurements (see Fig.10.6) underline that different species of both ROS and RNS in different charge states can be formed in the fragmentation of multiply-charged GABA.

Even though biochemists have been more interested in ROS due to their complexity and different roles in the living organisms,[294] RNS can be also important reagents for chemical damage. Furthermore, some of the produced fragments apart from their high reactivity, also do have catalytic properties in their neutral charge state like e.g. the vinylimine fragment, which is an important nitrogen donor.[295]

As we have seen, both doubly charged ROS and RNS (fragments with m/z = 26.5, 27.5, 28.5 amu) are produced in the fragmentation of $GABA^{z+}$ with different probability, depending on the charge state $z+$ of the GABA molecules and the charge localisation. Therefore, it is important to understand the different production ratios between ROS and RNS as a function of the initial charge state of GABA, i.e if they have been generated in double or triple electron capture processes. It is also important to remember that direct CE from both $GABA^{2+}$ and $GABA^{3+}$ is the dominant process and therefore $NH_2CH_2^+$, $CH_2CH_2^+$ and $COOH^+$ are the main fragments observed, but ∼10% of the fragmentation of $GABA^{2+}$ and ∼20% of $GABA^{3+}$ can lead to highly reactive ROS and RNS.

MD simulations indicate that the formation of doubly-charged ROS and RNS follows quite different pathways if we consider $GABA^{2+}$ or $GABA^{3+}$. In $GABA^{2+}$ the main mechanisms involve either sequential losses of neutral CO_2 and H_2 and in some instances H_2O or isomerisation processes in which doubly-charged diol isomers appear as important intermediates. As a consequence, only RNS^{2+} species are formed. It is important to stress that mechanisms leading to ROS would imply emission of neutral NH_3, a process that according to the MD simulations is hardly observed in $GABA^{2+}$ fragmentation.

In the case of $GABA^{3+}$, CE can also lead to the formation of stable doubly-charged fragments. If the $C_\beta - C_\gamma$ bond is broken, the double charge remains in the largest fragment and it contains the COOH group, being a precursor of doubly-charged ROS. However, if the $C_\alpha - C_{carboxilic}$ bond is broken, the double charge remains in the $NH_2CH_2CH_2CH_2$ group, and a precursor of doubly-charged RNS is obtained. Isomerisation processes in $GABA^{3+}$ mostly consists in H transfer forming a $-NH_3$ terminal group. Emission of NH_3^+ leads to ROS doubly-charged species. In $GABA^{3+}$, trajectories leading to the formation of charged diol derivatives are hardly obtained and therefore emission of water or CO does not occur.

Therefore, we propose different mechanisms for the production of the doubly charged moieties observed in the experiment, pointing out different charge localisations after double and triple electron captures:

$$GABA^{2+} \xrightarrow{\text{isomerisation}} [R - C(OH)_2]^{2+} \xrightarrow[-CO]{-H_2O} RNS^{2+}$$

$$GABA^{2+} \xrightarrow{\text{isomerisation}} [NH_3 - R]^{2+} \xrightarrow[-NH_3]{\text{not observed in the MD}} ROS^{2+}$$

$$GABA^{3+} \xrightarrow{CE} NH_2CH_2CH_2CH_2^{2+} + COOH^+ \rightarrow RNS^{2+}$$

$$GABA^{3+} \xrightarrow{CE} NH_2CH_2^+ + CH_2CH_2COOH^{2+} \rightarrow ROS^{2+}$$

$$GABA^{3+} \xrightarrow{\text{isomerisation}} [R - C(OH)_2]^{3+} \xrightarrow[-H_2O^+,-CO^+]{\text{not observed in the MD}} RNS^{2+}$$

$$GABA^{3+} \xrightarrow{\text{isomerisation}} [NH_3 - R]^{3+} \xrightarrow{-NH_3^+/-NH_4^+} ROS^{2+}$$

10.4 Conclusions

The combination of coincidence mass spectrometry with molecular dynamics simulations allows us to identify the main fragmentation channels from doubly-charged and triply-charged GABA generated in low energy collisions with Ar^{9+} ions. Although Coulomb explosion is the main mechanism for fragmentation of both $GABA^{2+}$ and $GABA^{3+}$, alternative channels can lead to the production of highly reactive RNS and ROS doubly-charged species. These species can be formed in processes involving Coulomb explosion, isomerisation or neutral fragment emission as a first step. In the case of $GABA^{2+}$, only RNS^{2+} species were identified, i.e. the charge remains located in the N-containing fragment, and the mechanism involves isomerisation plus neutral fragment emission. However, in the case of $GABA^{3+}$, the produced doubly-charged fragments correspond to both RNS and ROS, i.e. the double charge can be located either in the N-containing or the O-containing fragment. The main mechanism in this case starts with a CE, but alternative pathways involving isomerisation and emission of NH_3^+/NH_4^+ fragments can also lead to doubly charged ROS. The observation of a strong correlation between the production of reactive species and the total charge in GABA molecules offers interesting perspectives for the study of fragmentation mechanisms of multiply charged biomolecules and opens the door to the identification of new charged ROS and RNS that can be involved in chemical indirect effects in radiation damage.

Chapter 11

Results and Discussion

Stability and fragmentation of singly-charged *N*-acetylglycine

This chapter presents a combined experimental and theoretical study of the ionisation of N-acetylglycine molecules by 48 keV O^{6+} ions. We focus on the single ionisation channel of this interaction. In addition to the prompt fragmentation of the N-acetylglycine cation, we observe also the formation of metastable parent ions with lifetimes in the microsecond range. On the basis of density functional theory calculations, we assign these metastable ions to the diol tautomer of N-acetylglycine. In comparison to the simple amino acids, the tautomerisation rate is higher due to the presence of the peptide bond. The study of a simple biologically relevant molecule containing a peptide bond allows us to demonstrate how increasing the complexity of the structure influences the behaviour of the ionised molecule.

The results presented in this chapter has been published:
J. Kocisek, D. G. Piekarski, R. Delaunay, B. A. Huber, L. Adoui, F. Martín, M. Alcamí, P. Rousseau, A. Domaracka, J. Kopyra, and S. Díaz-Tendero, "*N*-acetylglycine cation tautomerisation enabled by the peptide bond", *The Journal of Physical Chemistry A* 119, 9581 (2015)

11.1 Introduction

Several experimental as well as theoretical studies concerning the ionisation of amino acids in the gas phase have been performed (see e.g.[92;176;209–211;238;251;296–300]). After single ionisation, the charge naturally localises on the amino group,[209] depending on the basicity of the side chain.[72] Independent of the ionisation method, isolated amino acids preferentially lose the carboxyl radical to form the $(M - COOH)^+$ ions followed by further fragmentation leading to the immonium cation $CH_2NH_2^+$.[210;251;298] An important question is - how

such fundamental knowledge obtained in the gas phase correlates with the processes occurring in living organisms? This question has been adressed in recent studies concerning the role of the environment in the ionisation of amino acids.[45;301;302] In this chapter, we complement these studies with a molecule containing a peptide bond in order to explore the behaviour of peptides and proteins. *N*-acetylglycine (aceturic acid, $C_4H_7NO_3$, see Fig. 11.1 a), a derivative of the simplest amino acid glycine (Fig. 11.1 b), is an ideal candidate for our purposes. It is large enough to evaluate the influence of a peptide bond and it is small enough to be thoroughly studied with theoretical simulations.

Recently, *N*-acetylglycine has been studied in experiments with low energy electrons (0–12 eV).[303] In this case, one electron attaches to the molecule to form a transient negative ion, which subsequently dissocites to form a fragment anion and neutral counterpart(s) - dissociative electron attachment (DEA). Electron attachment mainly results in the production of the $(M\text{-}H)^-$ anion, which is formed from the accommodation of the excess charge on the -COOH group followed by the loss of the H atom. A further prominent reaction proceeds via the cleavage of the C_α–N bond (c/z fragmentation) to form CH_2COO^-. This dissociation channel has been also observed in electron capture dissociation (ECD) experiments.[304;305] The striking observation in DEA was, however, the formation of the fragment ion due to the loss of a neutral water molecule. This reaction channel was not observed previously from DEA to single amino acids and appears to be characteristic for the molecule containing a peptide bond within the structure.[303;306–308] It should be noted here that the mechanism which governs the loss of the H_2O molecule is not resolved so far. The behaviour of the *N*-acetylglycine is different when the molecule is exposed to electrons with energy above the ionisation threshold. The electron ionisation (EI) mass spectrum resembles that of the oligopeptides, which is dominated by the production of b/y and $CH_2NH_2^+$ fragments.[309]

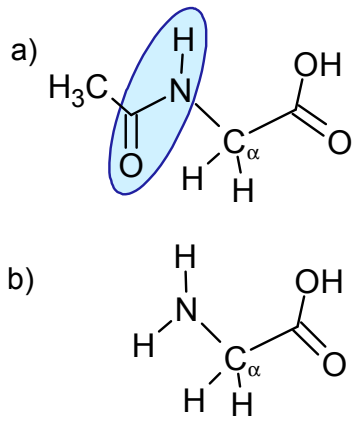


Figure 11.1: Sketch of a) *N*-acetylglycine molecule studied in the present work and b) glycine molecule. Peptide bond is highlighted and C_α carbon is labelled.

In the present study we use multiply charged ions to softly ionise the *N*-acetylglycine molecule in the gas phase and coincidence mass spectrometry techniques to explore the subsequent fragmentation. Quantum chemistry simulations allow us to infer the mechanisms behind the detection of the produced cationic fragments and give insight into chemical reactions that occur in positively-charged

molecules with a peptide bond in their structure. Special attention will be paid to the specific features of the metastable dehydration decay.

11.2 Methods

Experimental details

Experiments were carried out by means of the COLIMACON experimental setup.[92] *N*-acetylglycine molecules purchased from Sigma Aldrich company (99% purity) were evaporated in the molybdenum oven resistively heated at the temperature of 360 K. The temperature was monitored with a K-type thermocouple, allowing a sufficient vapour pressure while avoiding thermal decomposition of the molecule[303]. The vapour was passing through a long channel of the oven with a diameter of 2.5 mm. The formed effusive beam of molecules crossed the beam of O^{6+} ions from the ARIBE beam-line of the GANIL facility located in Caen, France. The kinetic energy of ions used in the present study was 48 keV. Cationic products of the interaction have been analysed by a Wiley-Mclaren TOF spectrometer operated in the coincidence regime (see Refs.[92;213] for details). The experiment was pulsed at a frequency of 4 kHz, the ion pulse length was 500 ns and the extraction pulse length was 6 μs. The extration field of 240V/cm is applied immediately after the ion bunch passage ($\approx$ 10 ns). In this configuration we can separate the isotopic contributions. The extraction region of the spectrometer was designed to reduce the initial energy spread of the ions due to the large interaction volume. Therefore, the length of extraction region is 11 cm. Such modification allows also to analyse the metastable decay of ions decaying on the μs timescale (see Ref.[310] for details).

Computational details

Quantum chemistry calculations were performed using the density functional theory (DFT). In particular geometry optimisation was carried out with the B3LYP functional, which combines the Becke's three parameter non-local hybrid exchange potential[114] with the non-local correlation functional of Lee, Yang and Parr.[215] This functional has been used in combination with the 6-311++G(d,p) basis set. This method has shown to provide an accurate description of the potential energy surface (PES) of cationic amino acids.[72;92;164;176;177;238]. Harmonic vibrational frequencies have been also evaluated at the same level to characterise minima and transition states in the PES and to compute the Zero Point Energy (ZPE) correction. In the computed electronic open shell structures (doublet spin multiplicity) we have checked that they do not present spin contamination; the highest value found for $< S^2 >$ is 0.7723 - i.e. it only differs in less than 3% with respect to the extact value. The calculations have been carried out with the Gaussian09 package.[195]

11.3 Results and discussion

Structure and isomerisation of neutral and cationic *N*-acetylglycine

We first present the theoretical results concerning the structure of neutral and singly-positively-charged *N*-acetylglycine together with the isomerisation between the different obtained structures. In order to generate the neutral conformers we have explored the PES by rotation of the dihedral angles and by considering all possible orientations of the carboxyl group. The geometry of the cationic isomers were obtained by geometry optimisation starting from the neutral ones. In the appendices we show the different conformers obtained for neutral *N*-acetylglycine including their relative stability; the notation "**c**" and "**t**" correspond to the *cis* and *trans* conformations with respect to the peptide bond. The obtained 15 isomers present a high degree of degeneracy (within less than 12 kcal mol^{-1}). The most stable structure **t6**, presents a *trans* conformation around the peptide bond and is stabilised by an intramolecular hydrogen bond between the carboxyl group and the H attached to the N atom (-N-H...O=C-). The computed isomerisation barriers between these conformers are shown at relatively low energy as well (see appendices), in particular the connection between the most stable *cis* and *trans* conformers (structures **c4** and **t6** respectively) requires only 6.9 kcal mol^{-1} from **t6**.

After ionisation of the neutral conformers followed by optimisation, we have obtained the cationic structures shown in the appendices. Again a high degree of degeneracy is observed for these conformers. The charge is mainly localised around the peptide bond and the intramolecular interactions are different with respect to the neutral structures. The most stable isomers also present a *trans* conformation and are connected with low energy transition states (see the low barriers connecting **t6**+, **t5**+ and **t10**+ in the appendices). These three structures are also connected with stable *cis* conformers (**c4**+, **c3**+ and **c7**+) through low energy transition states.

For the most stable neutral and cationic conformers (**t6** and **t6**+) we have performed an analysis of the wave function using the Quantum Theory of Atoms in Molecules.[156] We have found that intramolecular hydrogen bond only appears in the cationic conformer (in the bond critical point the electron density is 0.023 and the Laplacian is +0.101).

As we will show later, tautomerisation of the canonical structures towards a geminal diol structure is a very important process that explain the delayed fragmentation observed for cationic *N*-acetylglycine. Thus, we have also analysed different diol cationic conformers (see the appendices, where the notation **dt** or **dc** correspond to a diol isomer with *trans* or *cis* conformation around the peptide bond, respectively). All of them are more stable than any **t** or **c** cationic isomers. In addition, isomerisation between the cationic geminal diol isomers presents also low energy barriers that can be found below the most stable conformers with ca-

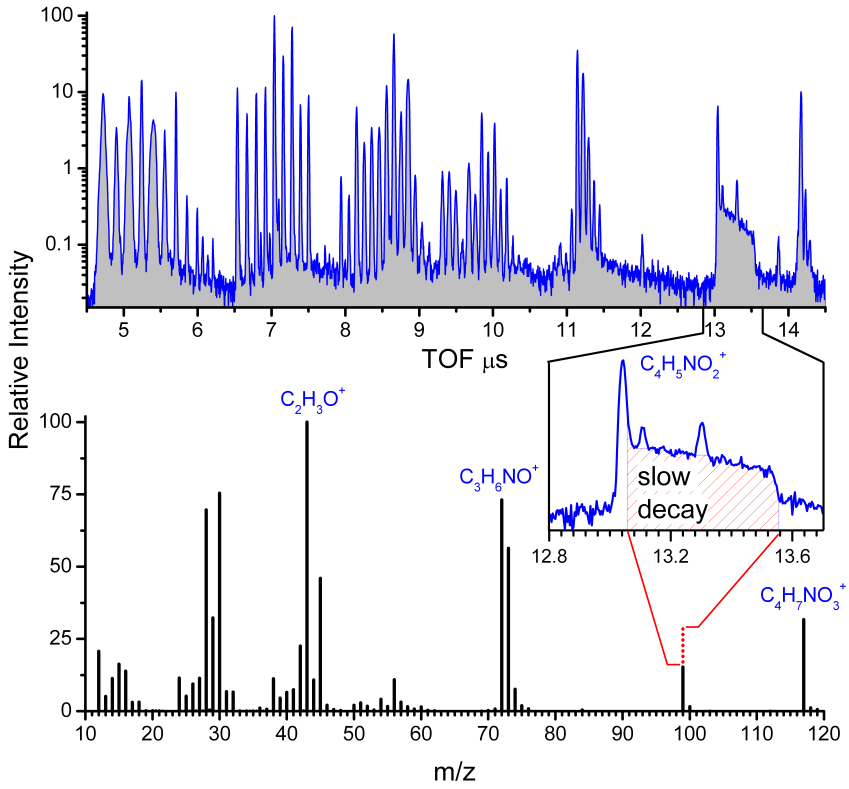


Figure 11.2: Single stop mass spectrum (MS) of product ions after interaction of 48 keV O^{6+} ions with *N*-acetylglycine molecule. Raw TOF spectrum including background contribution - top. Integrated peak intensities - bottom. The inset enlarges the TOF region around the $m/z = 99$ with two contributions to the observed $m/z = 99$ signal due to the fast and slow dissociation channel. In the integrated MS, the slow dissociation contribution is represented by red dots. After baseline subtraction, the peaks were integrated by a trapezoidal method. The integration was done for both, signal and background spectra. The integrated mass spectrum represents the difference between the signal and the background intensities individually determined for each m/z value.

nonical structure (see the appendices). The studied diol structures where the H attached to the carboxyl group does not come from a tautomerisation process but it was previously bonded to the N (structures **d** in the appendices) are much less stable than the canonical **t** and **c**, or the diol **dt** and **dc** cations.

Some of the considered isomerisation processes imply complex transition states involving simultaneous H transfer and COOH rotation (e.g. transition between **t10**+ and **dt11'**+ and between **c4**+ and **c7**+). They are one-step mechanisms leading to structures in which two H present different relative positions.

Prompt fragmentation

In our experiments we have studied interactions of O^{6+} ions with the N-acetylglycine molecule in the gas phase. The O^{6+} ions have been chosen for two reasons: (i) oxygen ions are promising projectiles for the cancer ion beam therapy,[311] (ii) O^{6+} ions easily allow to single ionize the molecule without large energy transfer[251]. The mass spectra of the produced cationic species were measured in the coincidence mode which enables the sorting of the spectra according to the number of detected ions per collision event. The mass spectrum with only one ion per collision event is called the single stop mass spectrum and it shows the singly-charged species produced in the collision and the stable multiply charged cations. The single stop mass spectrum and the time-of-flight raw data spectrum of N-acetylglycine are shown in Fig. 11.2. The highest signal is observed for the fragment ion at $m/z = 43$, that can be assigned either to CH_3CO^+ or to $OCNH^+$. The formation mechanism of the CH_3CO^+ ion (deacetylation) corresponds to the peptide bond cleavage. However, production of $OCNH^+$ requires rupture of two bonds: C_α−N and C−C (see Fig. 11.1a). In our previous study of the glycine molecule,[251] we have shown that the average energy deposited in the molecule by interaction with slow multiply-charged ions is rather low (of the order a few eVs). Hence, from the energetic point of view it is more likely that we generate CH_3CO^+ via the peptide bond cleavage. We also observe with high intensity in the spectrum a peak at $m/z = 72$. It is assigned to the loss of neutral COOH, a quite usual fragmentation channel in excited singly charged amino acids (see e.g.[92;210;211;251;298]). Our simulations (see Fig. 11.3a) show that the excitation energy required to dissociate into both paths (leading to $m/z = 43$ and 72) is 1.3 eV above the entrance channel (first vertical ionisation potential, 1st VIP). In glycine however, the most intense peak observed in the experiment ($m/z = 30$)[251] corresponds to a fragmentation channel that appears below the 1st VIP, $NH_2CH_2^+$ + COOH (see Fig. 11.3b); i.e. it does not require additional excitation energy.

Delayed fragmentation

The employed time-of-flight (TOF) spectrometer enables us to study the fragmentation of metastable ions at the μs timescale, which appears as a typical tail of the fragment ion signal towards higher TOF values.[310] Such behaviour is observed in the present work for the signal at $m/z = 99$ (see inset of Fig. 11.2). The residence time of the precursor ion in the extraction region of the mass spectrometer can be calculated on the basis of the ion TOF. In this way, the metastable parent ion can be identified from the maximum residence time which corresponds to the end of the tail. We assign the observed metastable decomposition reaction to the loss of neutral water:

$$C_4H_7NO_3^+ \rightarrow C_4H_5NO_2^+ + H_2O \tag{11.1}$$

The lifetime of the metastable parent ion can be calculated form the dependence of the $m/z = 99$ ion signal on the residence time in extraction region (see Fig. 11.4).

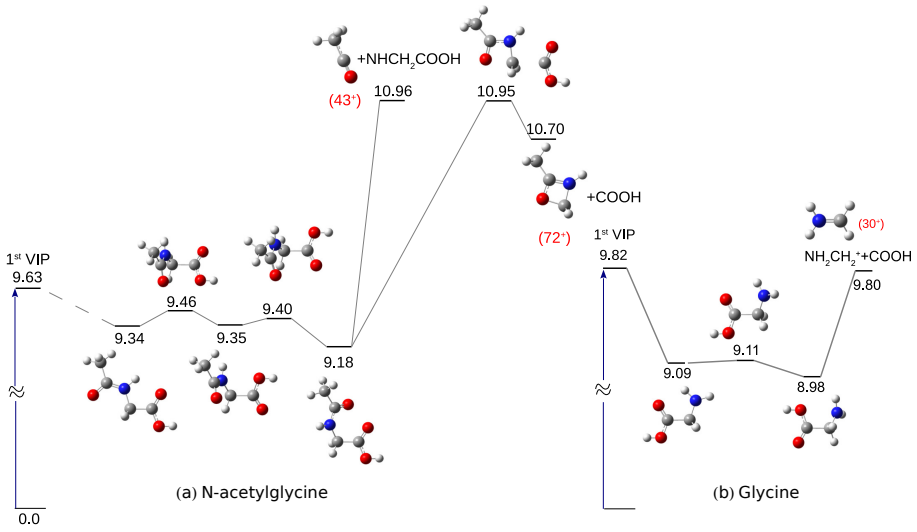


Figure 11.3: Energetic scheme of the lowest energy pathways for emission of COOH and peptide bond cleavage in *N*-acetylglycine (a) and for emission of COOH in glycine (b). 1st VIP represents the first vertical ionisation potential for the most stable neutral conformer (entrance channel). Stationary points (minima and transition states) of the potential energy surface (PES)are shown with relative energies given in eV at the DFT-B3LYP/6-311++G(d,p) level of theory including the zero point energy correction with respect to the most stable isomer of the corresponding neutral molecule.

In the present case, the lifetime of the metastable ions is in the microsecond range. Such long lifetime cannot be explained only by the hydrogen migration mechanism or other atomic rearrangements within the *N*-acetylglycine singly-charged cation since these processes should proceed on significantly shorter timescales[176;177;224]. Hence, an energetically stable molecular ion has to be formed.

In order to identify the structure of the stable molecular cation and its production, we have also computed the part of the potential energy surface of singly-positively-charged *N*-acetylglycine corresponding to the dissociation reaction Eq. (11.1). Figs. 11.5 and 11.6 present the lowest energy pathways found for this fragmentation channel from the most stable *cis* and *trans* conformers, respectively. They show an energetic scheme including the vertical ionisation and the further decomposition of these conformers. We have considered direct H_2O loss from the canonical forms of the ionised molecule, that proceeds via transition states from structures **c4**+ and **c3**+ in Fig. 11.5, and from structures **t6**+, and **t10**+ in Fig. 11.6. The ionised molecules can be also stabilised by tautomerisation into a diol form, leading to structures **dc4**+ in Fig. 11.5 and **dt11**+ in Fig. 11.6, that can evolve towards H_2O loss as well. The cationic diol is a well known energetically stable structure of glycine[49] which, analogously to our observation, also dissociate via

neutral water release. However, it has been shown that tautomerisation of cationic glycine has low probability due to high energy barriers and strong competition with other dissociative channels.[37;230] In agreement, our calculations show that both diol formation followed by H_2O emission and direct H_2O loss from the canonical form of cationic glycine present higher barriers than the one for the loss of neutral COOH (see Fig. 11.3b and Fig. 11.7). In contrast to that, metastable *N*-acetylglycine cations are observed with significant intensities in the present mass spectra (see Fig. 11.2). Hence, our experimental results indicate that the ionised molecule effectively tautomerises to the diol form with further emission of neutral H_2O.

For both molecules (*N*-acetylglycine and glycine) the lowest energy pathways of the two water release channels, directly from the canonical structure or via tautomerisation, look very similar. We observe (i) direct dissociation with an energy barrier of ∼1.7 eV (TSs at the right side of VIP in Figs. 11.5, 11.6 and 11.7) and (ii) a slow channel proceeding via the diol metastable ion (TSs at the left side of VIP in Figs. 11.5, 11.6 and 11.7) with a similar energy barrier for the water release. However, the rate determining step for the slow channel is the tautomerisation towards the diol form. This rate in the case of glycine should be much smaller due to the presence of the energetically more favourable channel $NH_2CH_2^+$ + COOH (Fig. 11.3b). In the case of *N*-acetylglycine all channels (tautomerisation, loss of H_2O, emission of COOH and peptide bond cleavage) appear at the same energy heights (∼1 eV above the entrance channel, see Figs. 11.3a and 11.5, 11.6).

It has been shown by Depke et al.[49] that the diol form of the glycine cation can be formed by subsequent hydrogen atom transfer from C_α to the amino group followed by hydrogen atom transfer from the amino to the carboxyl group. In such a way, the hydrogen atom can be transferred without significant structural changes. As we can see in the Fig. 11.6, this mechanism does not hold for the most stable *trans* isomer of cationic *N*-acetylglycine (structure **t10**+). In this case, the hydrogen atom on the nitrogen points towards the less favourable position before the tautomerisation, i.e. the geometrical arrangement prevents this mechanism. Then, the tautomerisation proceeds via direct hydrogen exchange between the C_α and carboxyl group, pointing out the importance of the peptide bond orientation. Such direct exchange could in principle be enabled by the favourable orientation of hydrogen in the TS due to the interaction between carboxyl and carbonyl groups. In order to evaluate the importance of this interaction, we performed calculations for the *N*-acetylglycine isomer with *cis* conformation around the peptide bond, which excludes such interaction. The calculated energetic barriers for water release have similar heights to that of the *trans* form (see Figs. 11.5 and 11.6). We can thus conclude that the different efficiency of the isomerisation process for the glycine and *N*-acetylglycine is caused purely by the effectiveness of the hydrogen transfer from the C_α to form the diol. Indeed, if we compare the charge distributions within glycine and *N*-acetylglycine after vertical ionisation (see Fig. 11.8), we can see that the main difference lies in the positive charge distribution on the carboxyl oxygen. In the case of glycine, the charge mainly localises on the nitrogen and partially on

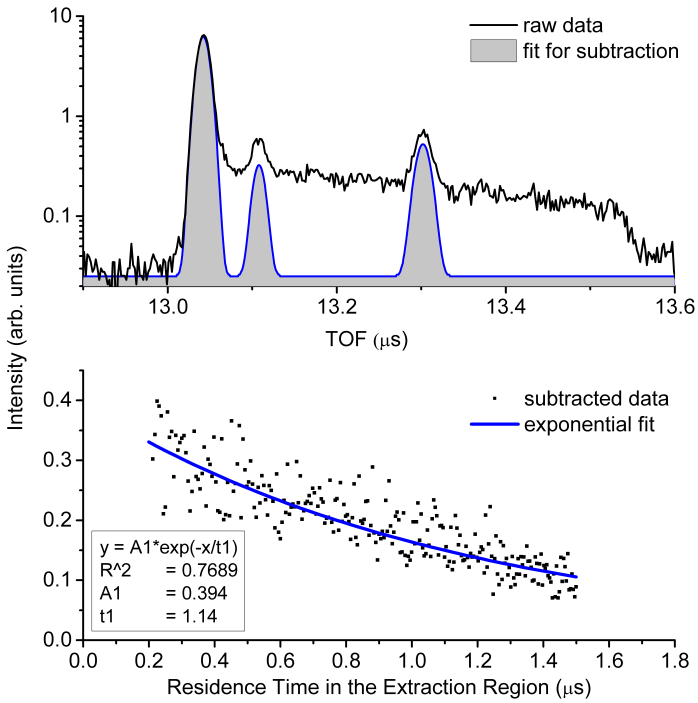


Figure 11.4: The top figure shows the TOF data around the $m/z = 99$ region including the raw data and the peaks corresponding to fast dissociation contribution (shadow area), which was subtracted for the fitting. The bottom graph shows the subtracted metastable decay data from the upper picture with the x axis converted into the residence time of the ion in the extraction region of the TOF spectrometer. The blue curve on the bottom figure is the exponential fit used to estimate the lifetime of the ions. The fit parameters are shown in the inset.

the carboxyl, whereas in the case of *N*-acetylglycine, the charge delocalises on the peptide bond while the charge on the carboxyl group remains nearly unchanged in comparison to the neutral molecule. Consequently, the carboxyl oxygen of *N*-acetylglycine cation has larger affinity for H than the one in glycine and effectively forms the diol structure. Furthermore, the electron density difference distributions also show that the $C_\alpha - C_{carboxyl}$ bond of glycine presents density depletion, i.e. the bond becomes weaker after ionisation and thus with larger probability to be broken with respect to the same bond in *N*-acetylglycine, which is polarised as a consequence of the adjacent peptide bond. This result strengthens the larger rate observed for COOH loss from glycine, and the changes observed in the electronic properties of *N*-acetylglycine due to the presence of the peptide bond.

The metastable ion yield is given by the reaction rates, as discussed before, and

the lifetime of the ion. The energy deposited in the system during the interaction with slow ions is in the range of a few electronvolts,[251] which is sufficient to overpass the energy barriers towards isomerisation as well as the water release. The lifetime of the metastable ion depends on its ability to redistribute the excitation energy. Since there are more degrees of freedom in the case of N-acetylglycine one can expect a longer lifetime of the metastable parent ion. This assumption is well confirmed by our simulations of isomerisation in the metastable diol cation structures. In our exploration of the potential energy surface, we have found 20 stable diol isomers (see the appendices) that appear at energies well below the barrier towards dehydration. Additionally, these structures are connected by low energy transition states (see the appendices). Therefore, after H migration to form one of the diol structures of N-acelylglycine, the molecular ion is trapped in a low energy valley of the potential energy surface with a large number of possible conformers. The ability to redistribute the energy and the confinement in the diol conformers provide a long lifetime of the excited cation towards H_2O emission, explaining the observed tail of the peak measured at $m/z = 99$.

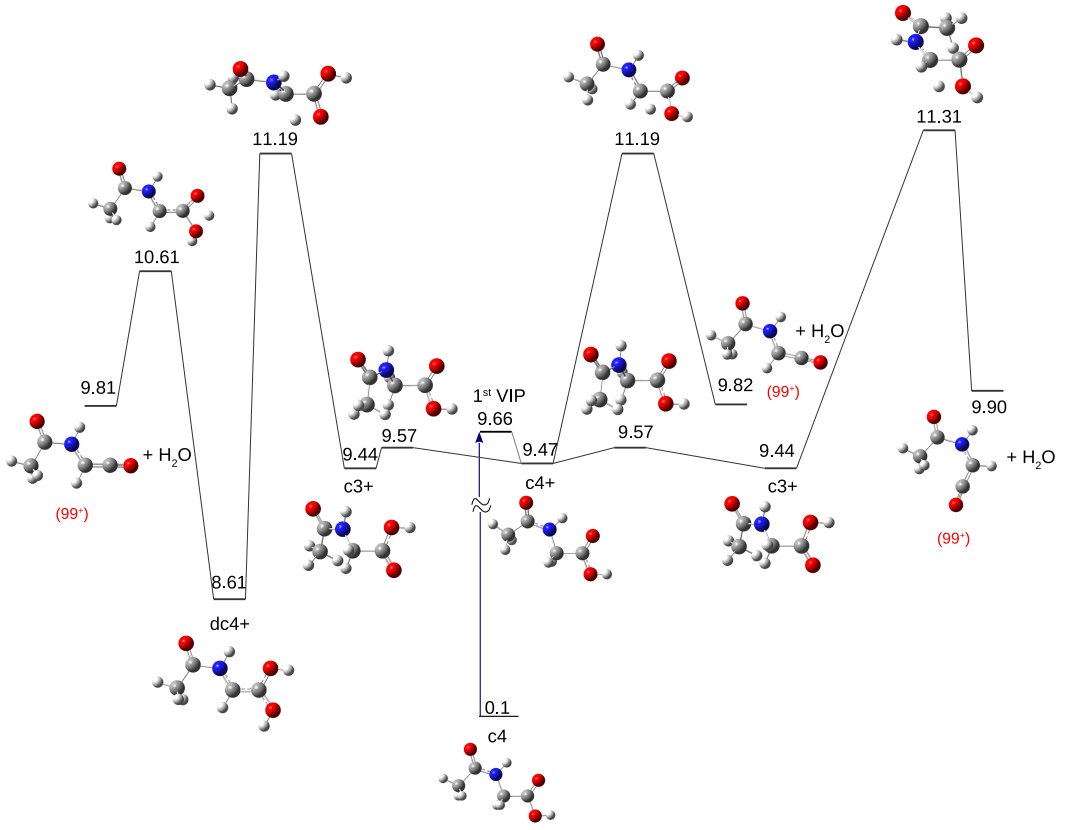


Figure 11.5: Fragmentation pathways for singly positively charged *N*-acetylglycine corresponding to *cis* ("**c**") geometry. Stationary points of the potential energy surface corresponding to direct and through geminal diol cation H_2O emission, right and left side of 1^{st} VIP, respectively, leading to charged fragments: $m/z = 99$. Relative energies are given in eV with respect to the most stable isomer of the neutral molecule (t6) at the DFT-B3LYP/6-311++G(d,p) level of theory including the zero point energy correction.

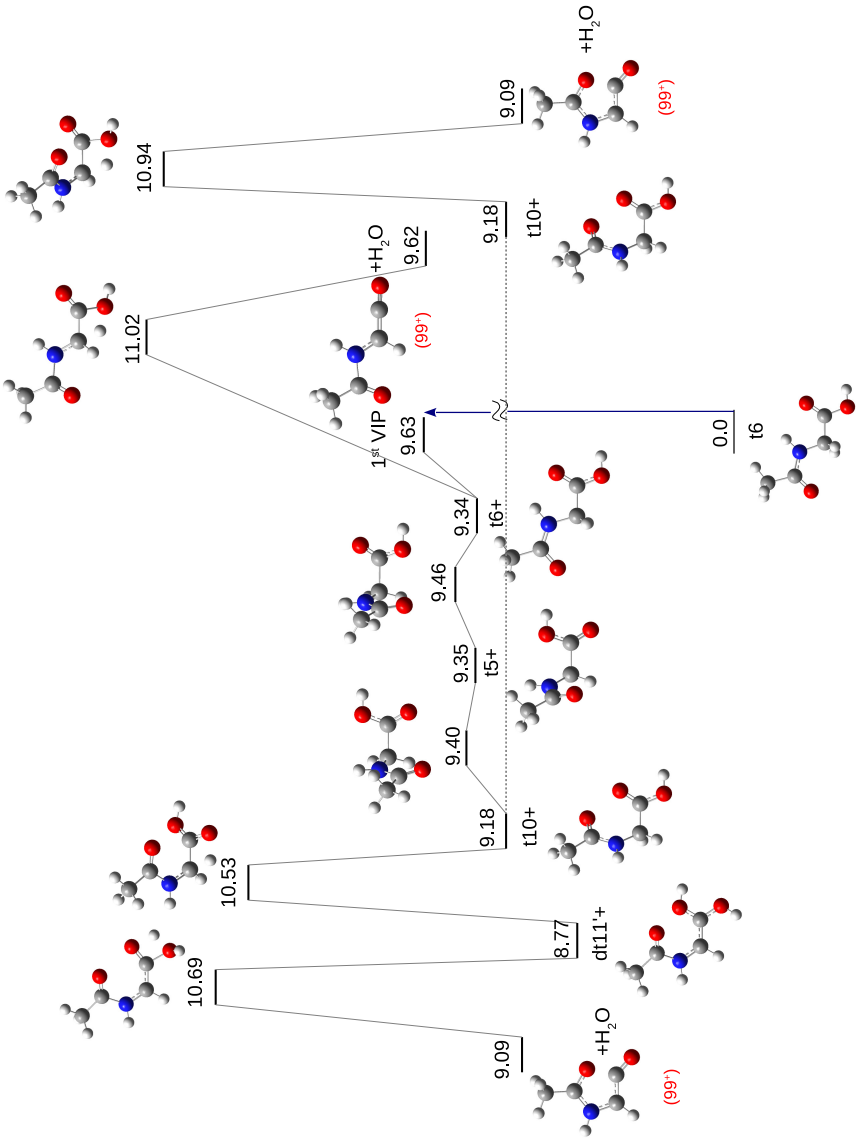


Figure 11.6: Fragmentation pathways for singly positively charged N-acetylglycine corresponding to *trans* ("**t**") geometry. Stationary points of the potential energy surface corresponding to direct and through geminal diol cation H_2O emission, right and left side of 1^{st} VIP, respectively, leading to charged fragments: $m/z = 99$. Relative energies are given in eV with respect to the most stable isomer of the neutral molecule (t6) at the DFT-B3LYP/6-311++G(d,p) level of theory including the zero point energy correction.

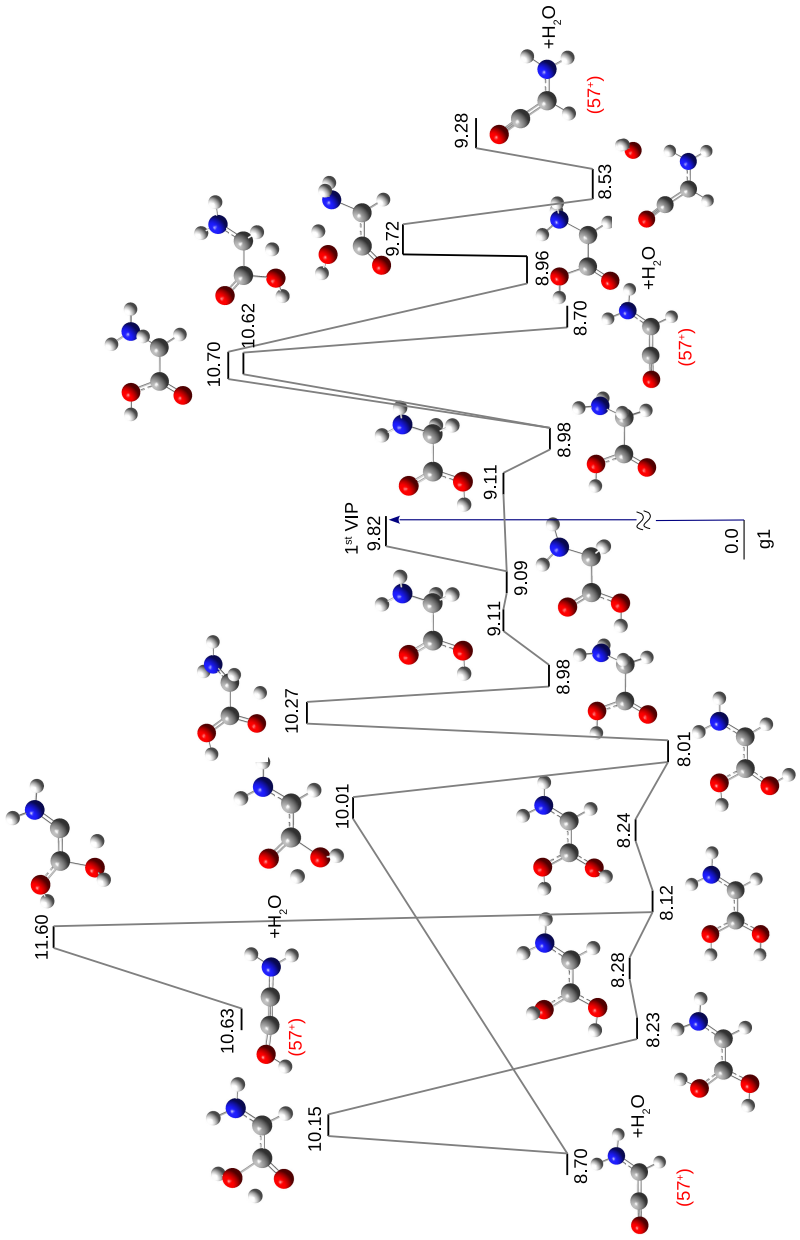


Figure 11.7: Fragmentation pathways for singly positively charged glycine. Stationary points of the potential energy surface corresponding to direct and through geminal diol cation H_2O emission, right and left side of 1^{st} VIP, respectively, leading to charged fragments: $m/z = 57$. Relative energies are given in eV with respect to the most stable isomer of the neutral molecule (g1) at the DFT-B3LYP/6-311++G(d,p) level of theory including the zero point energy correction.

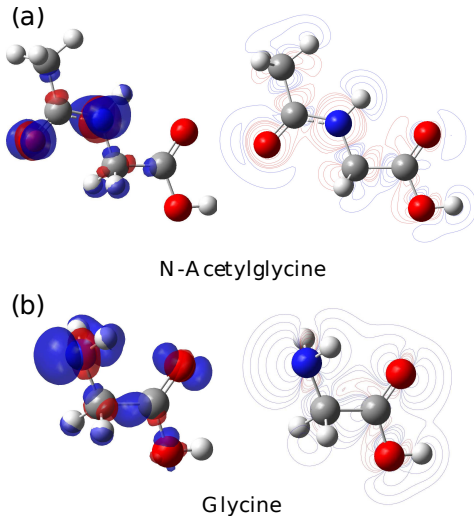


Figure 11.8: Electron density difference between neutral and ionised (a) N-acetylglycine and (b) glycine in a Franck-Condon transition; electron density depletion/accumulation of the electron density is shown in blue/red. Left panels 3D density difference at isovalues +/- 0.008 e^-/au^3.). Right panels present contour plots in the plane of the peptide bond (a) and of the molecular symmetry plane (b).

11.4 Conclusions

After ionisation and excitation of N-acetylglycine molecules by slow (48 keV) O^{6+} ions in the gas phase, we have observed fast fragmentation dominated by the peptide bond dissociation and neutral COOH loss. Surprisingly, we have also detected parent molecular cations with stability on the microsecond timescale. The long lifetime of this cation is assigned to tautomerisation processes leading to diol structures of remarkable stability. The observed tautomerisation rate is much higher in comparison to the glycine cation due to the effective charge delocalisation on the peptide bond and the different energetic barriers, which influence the dissociation dynamics. We have shown that in comparison to amino acids the behaviour of the produced ion is different, since the typical COOH loss competes with other channels, mainly the peptide bond rupture. Additionally, the reduced charge localisation on the carboxyl oxygen increases the mobility of the C_α hydrogen atoms and enables an effective tautomerisation towards the diol formation. The differences observed in the studied system with respect to simple amino acids, shed light on the mechanisms behind the fragmentation of complex molecular cations containing peptide bonds.

Chapter 12

Results and Discussion

Structure and stability of neutral clusters of β-alanine

This chapter presents a theoretical study of neutral clusters of $\beta-$alanine molecules in the gas phase, $(\beta - ala)_n$ $n \leq 5$. Classical molecular dynamics simulations carried out with different internal excitation energies provide information on the clusters' formation and their thermal decomposition limits. We also present an assessment study performed with different families of density functionals using the dimer, $(\beta - ala)_2$, as a benchmark system. The M06-2X functional provides the best agreement in geometries and relative energies in comparison with the reference values computed with the MP2 and CCSD(T) methods. The structure, stability, dissociation energies and vertical ionisation potentials of the studied clusters have been investigated with this functional in combination with the 6-311++G(d,p) basis set. An exhaustive analysis of the intermolecular interactions is also presented. These results provide new insights into the stability, interaction nature and formation mechanisms in clusters of amino acids in the gas phase.

The results presented in this chapter has been published:
D. G. Piekarski and Sergio Díaz-Tendero, Structure and stability of clusters of β-alanine in the gas phase: importance of the nature of intermolecular interactions *Physical Chemistry Chemical Physics*, 19, 5465 (2017).

12.1 Introduction

Clusters of β-alanine molecules are interesting to study from the chemical and biological point of view. Since they do not present covalent bonds between the residues, only the non-covalent interactions such as hydrogen bonds (HBs) and van der Waals forces stabilise them; The chemical behaviour (reactivity, structure, stability, etc.) and biological activity (highly specific recognition, transport,

regulation, etc.) of biomolecules are driven by these interactions.[312–314] It seems reasonable to assume that β-alanine molecules linked into a cluster present greater sensitivity than isolated molecules to key kinetic events, such as binding, insertion and dimerisation similarly to peptides.[315] In addition, these clusters could be considered as a relevant system for supramolecular chemistry, where intricate coalescence interactions within the cluster provide the formation of molecular cages. These complex structures are ideal systems to be analysed in detail in order to characterise the intermolecular interactions and understand the stability of such cages. Such supramolecular structures are stabilised with non-covalent forces and their study can shed light on intermolecular hydrogen bonding processes, where two or more residues interact as a target-receptor centers with donor-acceptor atoms for a hydrogen atom (D−H− − −A). A reliable study of the nature of these interactions provides also additional information to understand sizes, shapes and affinities in the formation of clusters of amino acids[316] and other biomolecules.[317–319] The knowledge of stereospecific intermolecular hydrogen bonding could provide specific information about cell permeability and other physicochemical properties.[320] Part of the present work focuses on disentangling the nature of intermolecular interactions between β-alanine molecules inside the clusters and has a major interest also for chemical-biology, where they play a crucial role for the proteins recognition and aggregation.[321] The findings may also contribute to a more general understanding of the role of weak interactions in protein folding. Moreover, β-alanine clusters could be considered as the simplest system to study amino acids all-side-chain interactions via intermolecular hydrogen linkers creating unnatural and stable cages, similarly to binding proteins domain, through intermolecular β-strand interactions.[322;323] Thus, analysing the structure and the stability of clusters of β-alanine molecules reveals the role played by forces involved in similar chemical models like e.g. complex protein structures, where the intramolecular self-assembly and intermolecular β-sheet interactions are crucial.[324] Though solvent effects should be included to mimic the environment of proteins, in this chapter we focus on the role that these interactions play in stabilising the clusters in the gas phase.

The structure and stability of biomolecules in the gas phase and the strength of such inter- and intramolecular weak forces are challenging to predict. Combining spectroscopy with theoretical simulations, folding of isolated neutral tetrapeptides[325] and three-dimensional structures of peptides[326] have been successfully unraveled. Unexpected selective formation of dimers between similar biomolecules has been also a subject of investigation.[327] Long-range interactions towards dimerisation processes in artificial β-sheets has been pointed out as an important effect.[328] The conformational analysis has become a fundamental tool to determine the size and structure of weakly bonded molecular clusters.[329–331] For instance, it has been demonstrated that water hexamer cage, stabilised by HBs is the global minimum.[332] The structure and the stability of clusters of amino acids (in this case clusters of neutral β-alanine) are also key aspects to understand physicochemical phenomena related to dynamics of such complex systems.

Experimentally, intermolecular interactions are difficult to measure directly,[333]

and most experiments rely on indirect analysis.[334–336] It has been proposed experimentally[337] that oligomers constructed from β-amino acids avoid interaction through HBs between the nearest neighbors of the amide groups along the backbone. Thus, they become an attractive "material" to build the foldamers since abstention of attractive interactions with the nearest neighbors appears to be important for acceptance of compact secondary structures. On the other hand, ion collision experiments have revealed that amino acids embedded in a molecular cluster, are much likely to keep their structure without undergoing fragmentation due to the protective effect of the environment, showing the importance of the intermolecular interactions.[45] In this chapter we also evaluate which are the main contributions behind this behaviour providing useful information such as dissociation energies and ionisation potentials. Likewise, our study facilitates understanding of the origin of the discrepancies between different theoretical studies and experimental results in the stabilisation energies in clusters of amino acids.[338] In light of an exhaustive analysis of crystal structure databases, the intermolecular HBs are found to be involved in the formation of up to six-membered rings.[339] Our unexpected finding is already observed for dimers where the intermolecular weak interactions lead to the formation of a twelve-membered ring. For bigger clusters several many-member rings are observed and more complicated 3D structures, similar to cages, were found. We analyse the importance of these rings and we show cooperative effects within them.

Clusters of β-alanine molecules represent a challenging system to investigate theoretically. Molecular dynamics at a high level of theory, such as *ab initio* molecular dynamics, are unfordable for several reasons. Firstly, the size of the cluster increases the amount of computer time in each time step. Secondly, to properly populate the conformational space the simulation time should be much longer than for single molecules due to the huge number of degrees of freedom of the system. Thirdly, in systems with weak interactions it is difficult to achieve convergence in energies and forces. Therefore, molecular clusters are limited to less demanding methods as semi empirical or classical molecular dynamics simulations. Static simulations on molecular clusters can be carried out with more accurate *ab initio* methods up to certain sizes. A correct description of weakly bonded systems presenting short-range correlation and long-range dispersion forces is not a trivial task and also involves a further theoretical effort.[117;340–342] In this chapter we present a computational strategy that combine molecular dynamics simulations with density functional theory calculations to study molecular clusters. We implement this strategy in clusters of β-alanine molecules with sizes up to five molecules. We also present a benchmark study that provides useful information on the correct choice of the functional in order to accurately describe these systems at the density functional theory (DFT) level, based on energetic and geometrical criteria. A systematic evaluation of the intermolecular interactions allows us to shed light on the importance of their nature in the structure, stability and formation of weakly bonded clusters of biologically relevant molecules. We also present dissociation and vertical ionisation energies.

12.2 Computational Methods

Our study was performed for clusters of the lineal amino acid β-alanine, including up to five molecules, and it combined molecular dynamics simulations (MD) with density functional theory (DFT) calculations. MD simulations provide information of the formation of the clusters as a function of temperature and is a very useful tool to populate many stable isomers; it could be also easily extended to bigger systems. Benchmarking of a few families of functionals was carried out by comparing the obtained geometries and energies with second order Møller-Plesset perturbation theory (MP2) results. The HBs were characterised with the quantum theory of atoms-in-molecules (QT-AIM),[156] the natural bond orbital (NBO) theory,[155;343] and the Non-Covalent Interactions (NCI) visualisation tools.[344;345]

Molecular Dynamics

The structure and stability of neutral clusters of β-alanine as a function of the temperature were determined with molecular dynamics based on the general amber force field, GAFF,[147] with RESP charge,[148] using the Assisted Model Building and Energy Refinement program[136–139] (AMBER12-5.0 package). The performace of the GAFF force field was checked on nucleic acid base pairs in gas phase, showing good agreement with *ab initio* calculations (MP2/6-31G*) for minimised structures and intermolecular energies.[147] MD using GAFF was then employed to explore a wide range of the potential energy surface (PES). To this, the system was slowly heated until it reached the target temperature and then it was propagated during a given time. The Berendsen thermostat[346] was used to ensure a constant temperature for the heating stage with a temperature coupling parameter of 10.0 ps in the microcanonical (NVE) ensemble. The Langevin thermostat[347] was used for the propagation dynamics with a collision frequency of 5.0 ps^{-1} in the canonical (NVT) ensemble (see convergence information in the appendices). The simulations were carried out using the velocity Verlet algorithm to integrate the Newtonian equations of motion with a time step of 0.2 fs, and a total simulation time of $3.6x10^3$ ps (heating stage of $8x10^2$ ps and propagation stage of $2.8x10^3$ ps). This propagation time should be long enough to populate many stable conformers of β-alanine clusters according to previous studies of small β-peptides (with 4-6 monomers).[348] We performed 10 trajectories for different temperature values considered: T= 50, 100, 150, 200, 250 , 300, 350, 360, 375, 400 and 423K. The critical temperature that maintains the cluster stability within the simulation time was found to be 300K. Above that temperature thermal decomposition of the clusters was observed.

The final geometry of each trajectory was taken as an initial guess for further optimisation with the DFT methods (several structures were obtained at each temperature, see final geometries in the appendices). For all the MD trajectories with dimers, after the propagation stage, we have extended the simulations by performing cool down dynamics, descending the temperature until 10K. We have checked that in all these cases the relative orientation between the monomers and

the connectivity remains the same as in the last step of the propagation. It is thus reasonable to think that geometry optimisation of these structures will lead the same conformers.

Additionally, the conformational search in the case of the dimer $(\beta - ala)_2$ was extended using the MMFF94[349] force field and a Monte Carlo algorithm as implemented in the Spartan14 package.[350;351] The MMFF94 force field was developed to describe a wide range of organic chemistry compounds and reactions. Thus, with this search we have checked the performance of two force fields developed for different objectives. The 45 most stable conformations found with MMFF94 were chosen among the 10000 simulations for further optimisation at the DFT level. As a result of the two conformational searches (molecular dynamics and Monte Carlo), a total of 19 different conformers of dimer were chosen for the benchmark study (see optimised geometries in the appendices). The conformers obtained using the Monte Carlo method appeared at higher energy (less stable structures) than those found with the AMBER MD method, after optimisation with DFT.

DFT calculations

Quantum chemistry calculations were carried out using the density functional theory (DFT). In particular, geometry optimisations were performed using different families of DFT functionals in combination with the 6-311++G(d,p) basis set. The combination of this basis set with some DFT functionals has shown to provide an accurate description of the PES of cationic amino acids.[72;92;164;176;177;238] Harmonic vibrational frequencies were also evaluated at the same level to characterise minima and to compute the Zero Point Energy (ZPE) correction. All DFT calculations were performed using the Gaussian09 program.[195]

We also studied electron density and bonding properties of the most stable structures using the quantum theory of atoms-in-molecules (QT-AIM) developed by Bader.[156]

Visualisation of the Non-Covalent Interactions (NCI) was carried out with the NCI-plot3.0 code,[344;345] through the analysis of the reduced density gradient.

In addition, we examined intermolecular interactions using natural bond orbitals computed with the NBO program.[352] The estimated strengths of the electron donor-acceptor interactions have been computed from the NBO analysis. A second-order perturbative analysis was used to quantify all possible intermolecular interactions between the donor residue and the acceptor one. The corresponding computed stabilisation energy ($\Delta E_{ij}^{(2)}$) is defined as:

$$\Delta E_{ij}^{(2)} = -q_i \frac{|F_{ij}|^2}{\epsilon_j^{(NL)} - \epsilon_i^{(L)}} \tag{12.1}$$

, where q_i is the occupancy of the donor orbital, F_{ij} is the off-diagonal NBO Fock matrix element and $\epsilon_i^{(L)}, \epsilon_j^{(NL)}$ are the donor and acceptor orbital energies, respectively. The HB stabilisation energy is defined as the sum of all donor acceptor

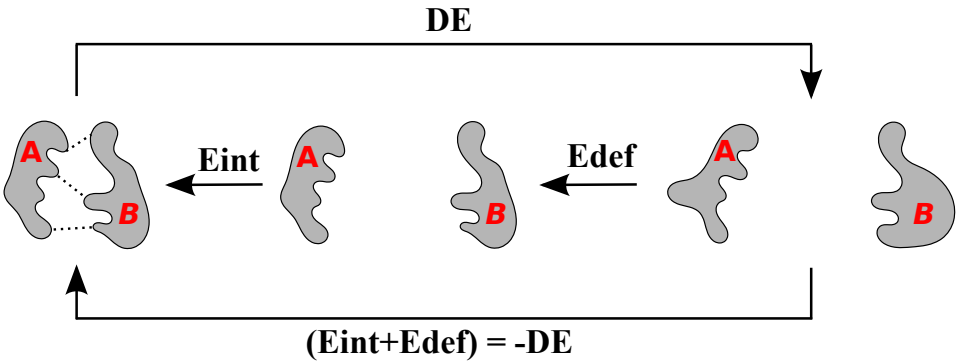


Figure 12.1: Schematic picture showing the relations between E_{int}, E_{def} and DE in the dissociation of neutral β-alanine clusters.

orbital interactions for a particular HB. Therefore, we compute the stabilisation energy of each HB (ϵ_{NBO}) as: $\epsilon_{NBO} = \sum(\Delta E_{ij}^{(2)})$, where the sum runs over all interactions in which orbitals of D, H and A atoms are involved (see detailed information in appendices). The total stabilisation energy (all the stabilisation energies within the cluster, E_{NBO}) is defined as a sum of all HBs stabilisation energies: $E_{NBO} = \sum(\epsilon_{NBO})$, where the sum runs over all HBs in the cluster.

Dissociation energies

Our methodology includes the zero point energy (ZPE) correction together with elimination of the basis set superposition error (BSSE) with the counterpoise correction (CP) [126] to provide accurate values of dissociation energies. For each dissociation channel we distinguish three different physical quantities: interaction energy (E_{int}), deformation energy (E_{def}) and dissociation energy (DE). The schematic picture showing the relation between these quantities is presented in Fig. C.8.
The interaction energy is defined as the attraction energy between the residues keeping the molecules linked in the cluster. Thus, the interaction energy between fragments **A** and **B** within the cluster **AB** is computed as the difference between the electronic energy of the optimised cluster obtained with the cluster basis set (E_{AB}^{AB}) and the electronic energy of the residues obtained with the corresponding residue basis set E_A^A and E_B^B, taking the geometry of **A** and **B** as in the optimised cluster **AB**:

$$E_{int} = E_{AB_{opt}}^{AB} - (E_A^A + E_B^B) \tag{12.2}$$

, where the *superscripts* indicate the basis set, and *opt* index refers to the optimised structure. The interaction energy corrected with the Counterpoise Method (CP) [353] is computed as:

$$E_{int}^{CP} = E_{AB_{opt}}^{AB} - (E_A^{AB} + E_B^{AB}) \tag{12.3}$$

The deformation energy is defined as the energy needed to modify the geometry of the fragments from their equilibrium geometry when they are isolated ($E^A_{A_{opt}}$ and $E^B_{B_{opt}}$) to the geometry that they have in the cluster (E^A_A and E^B_B). Then, the deformation energy (E_{def}) is computed as:

$$E_{def} = (E^A_A - E^A_{A_{opt}}) + (E^B_B - E^B_{B_{opt}}) \tag{12.4}$$

Dissociation energy (DE) is defined as the energy required to separate the fragments by bringing them to their optimised and isolated structures. Then the dissociation energy computed with the interaction energy including CP correction eq. (12.3) and the deformation energy eq. (12.4) becomes:

$$\begin{aligned} DE^{CP} &= -E^{CP}_{int} - E_{def} \\ &= -\left[E^{AB}_{AB_{opt}} - (E^{AB}_A - E^{AB}_B)\right] - \left[(E^A_A - E^A_{A_{opt}}) + (E^B_B - E^B_{B_{opt}})\right] \end{aligned} \tag{12.5}$$

If we correct the dissociation energy including the Zero Point Energy (ZPE)[354] of the optimised structures, we obtain:

$$DE^{ZPE} = \left[(E^A_{A_{opt}} + ZPE_A) + (E^B_{B_{opt}} + ZPE_B)\right] - \left[(E^{AB}_{AB_{opt}} + ZPE_{AB})\right] \tag{12.6}$$

The fully corrected dissociation energy (DE^{CP+ZPE}) is obtained including both CP and ZPE corrections:

$$\begin{aligned} DE^{CP+ZPE} &= -\left\{(E^{AB}_{AB_{opt}} + ZPE_{AB}) - (E^{AB}_A - E^{AB}_B)\right\} \\ &\quad -\left\{\left[E^A_A - (E^A_{A_{opt}} + ZPE_A)\right] + \left[E^B_B - (E^B_{B_{opt}} + ZPE_B)\right]\right\} \end{aligned} \tag{12.7}$$

Benchmarking

In order to find a suitable computational tool for carrying out reliable electronic structure calculations for clusters of β-alanine at a relatively low computational cost, a benchmark study was performed for β-alanine dimers. In our study we included different families of functionals: hybrid (namely B3LYP,[114;215;356] M06[117] and its variation M06-2X[117]), hybrid-meta (MPWB1K[357]) and dispersion corrected (B97D[341]) functionals in combination with the 6-311++G(d,p) basis set. The MP2 method with the same basis set was taken as a reference for the chosen model system $(\beta - ala)_2$. 19 conformers of $(\beta - ala)_2$ were considered (see the structures in the appendices).

The relative energies (including ZPE corrections), of the 19 dimers after geometry optimisation with the DFT methods and with MP2 are presented in the appendices. A very good agreement was found between the M06-2X functional and the MP2 method. The quality of the linear correlation between the relative energy calculated at MP2 and DFT levels is evaluated as the square of the Pearson

correlation coefficient (R^2). Almost complete correlation (R^2=0.99) is observed for the M06-2X functional. In addition, this functional gives the closest description in the relative stability order of the most stable conformers (see appendix F). Since the MP2 method was pointed out to be sometimes not reliable for relative energies of neutral gas phase peptides,[358] we have also performed single point energy calculations at the CCSD(T) level over the geometry optimised at the MP2 for the most stable dimers. The results confirm the quality of the M06-2X functional to compute the relative energies of the clusters (see appendix F).

We compare geometrical changes between the MP2 and the DFT results by calculating the root mean square deviation ($RMSD$) for the obtained structures, defined as:

$$RMSD = \sqrt{\frac{\sum_{i=1}^{N} \left(r_i\left(MP2\right) - r_i\left(DFT\right)\right)^2}{N}} \tag{12.8}$$

where N is the number of atoms whose position is compared, and r_i is the position of i atom. The B97D and M06-2X functionals accurately describe the geometry of almost all dimers in comparison with MP2. The average $RMSD$ of both of them with respect to MP2 is the same, 0.14 Å. On the other hand, only the M06-2X functional reproduces the geometry obtained at MP2 for conformer (2.9).

A comparative study of the QT-AIM results using all above-mentioned functionals and the MP2 method for the most stable dimer is presented in the appendices. In particular, the electron density and the Laplacian of the electron density in two bond critical points were evaluated (BCP1 and BCP2 – see bonds labeled 1 and 2 in Fig. 12.2(a)); These points correspond to intermolecular interactions in the most stable dimer structure. B3LYP and M06-2X functionals give exact values in comparison with MP2 results.

Overall, the results obtained with the empirically dispersion-corrected funcitonal B97-D and the hybrid functional with a significant Hartree-Fock exchange term M06-2X are in good agreement with high-level *ab initio* methods. These functionals have been previously shown good performance in the computation of gas-phase peptide chains.[358] It is thus remarkable the importance of a proper description of weak interactions such as dispersion forces and hydrogen bonds.

Thus, the M06-2X method was chosen because it adequately reproduces both the relative energies and geometries of the considered 19 isomers of $(\beta - ala)_2$, together with electron properties in the BCPs for the most stable dimer.

12.3 Results

Using the computational strategy described above we populated a large number of structures for each cluster size (dimer, trimer, tetramer and pentamer) and later, the M06-2X functional in combination with the 6-311++G(d,p) basis set was used to optimise the geometry and to evaluate the relative stability of the numerous isomers considered. We can see that the relative stability is strongly

linked to the structure via specific hydrogen bonds, HBs, for each cluster size (see all studied structures in the appendix F). For the sake of clarity and conciseness, only the most stable conformer for each cluster size is presented and analysed in the manuscript (see Fig. 12.2). Firstly, the dissociation energies of the clusters are presented. Then the interpretation of the obtained results is carried out on the basis of intermolecular stabilisation effects. To this, we present a thorough study based on wave function analysis techniques: QT-AIM and NBO theory. Finally, the last subsection presents a general overview of the stability of all found conformers of each cluster size against ionisation.

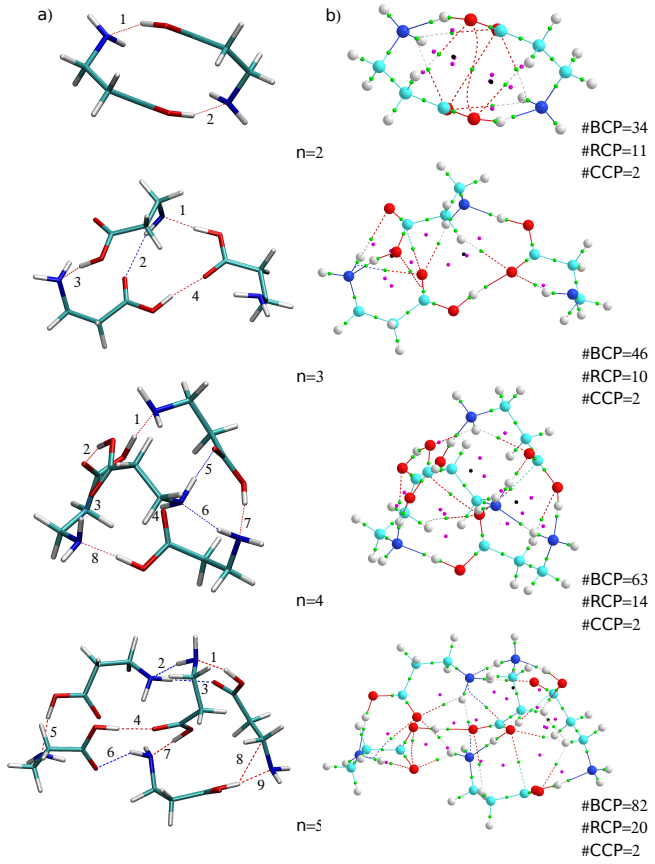


Figure 12.2: Structure of the most stable neutral clusters of $(\beta-ala)_n$ $n \leq 5$, calculated at the DFT-M06-2X/6-311++G(d,p) level of theory. Left panel: dashed lines shows HBs with the cut-off distance between donor (D) and acceptor (A) atom ≤ 3.2 Å and $D-\mathrm{H}-A$ angle ≤ 50 degree;[355] the colours of the HBs underline the donor of the hydrogen atom D: oxigen in red, nitrogen in blue and carbon in turquoise. Indexes label the corresponding HB. Right panel: topology of the electron density using QT-AIM analysis[156]; dashed black and solid gray lines connect the bond critical points (BCP, green dots) and ring critical points (RCP, magenta dots), respectively; black dots represents cage critical points (CCP).

Table 12.1: Interaction (E_{int}^{CP}), deformation (E_{def}) and dissociation (DE) energies in kcal·mol^{-1}

$(\beta-ala)_n \longrightarrow$ n	$(\beta-ala)_{n-m}+$ $n-m$	$(\beta-ala)_m$ m	E_{int}	E_{int}^{CP}	E_{def}	$BSSE$	ZPE	DE^{CP}	DE^{ZPE}	DE^{CP+ZPE}
2	1	1	-27.93	-25.83	5.05	2.10	-2.00	20.77	20.87	18.77
3	2	1	-40.46	-37.85	26.91	2.61	-0.22	10.94	13.32	10.72
			-29.48	-27.50	15.93	1.98	-0.22	11.57	13.32	11.34
			-32.06	-30.30	18.51	1.76	-0.22	11.79	13.32	11.56
4	3	1	-50.69	-47.53	27.04	3.16	-2.00	20.48	21.65	18.49
			-27.69	-25.70	4.05	2.00	-2.00	21.65	21.65	19.65
			-44.90	-42.05	21.25	2.85	-2.00	20.80	21.65	18.80
			-48.91	-45.65	25.26	3.26	-2.00	20.39	21.65	18.39
	2	2	-57.65	-53.82	43.33	3.83	-0.22	10.49	14.10	10.27
			-56.89	-53.00	42.57	3.90	-0.22	10.42	14.10	10.20
			-57.82	-54.26	43.50	3.56	-0.22	10.76	14.10	10.54
5	4	1	-35.58	-33.33	18.21	2.24	-1.50	15.13	15.87	13.63
			-42.97	-40.15	25.60	2.82	-1.50	14.55	15.87	13.05
			-54.07	-50.91	36.70	3.15	-1.50	14.22	15.87	12.72
			-64.39	-60.38	47.02	4.01	-1.50	13.36	15.87	11.86
			-35.02	-32.47	17.65	2.55	-1.50	14.82	15.87	13.32
	3	2	-74.05	-69.42	55.91	4.64	-1.49	13.50	16.65	12.01
			-78.44	-73.46	60.30	4.99	-1.49	13.16	16.65	11.67
			-66.71	-62.33	48.57	4.39	-1.49	13.76	16.65	12.27
			-59.75	-55.86	41.61	3.90	-1.49	14.24	16.65	12.75
			-34.51	-32.46	16.37	2.06	-1.49	16.09	16.65	14.60
			-75.50	-70.26	57.36	5.24	-1.49	12.90	16.65	11.41
			-76.56	-71.47	58.42	5.09	-1.49	13.05	16.65	11.56
			-81.33	-75.95	63.19	5.38	-1.49	12.76	16.65	11.27
			-65.47	-60.60	47.33	4.87	-1.49	13.27	16.65	11.78
			-65.87	-62.12	47.73	3.75	-1.49	14.39	16.65	12.90

Dissociation energies

The strength of the interactions in the clusters has been first evaluated in terms of dissociation energies. To compute them we have considered the most stable structures of the dimer, trimer, tetramer and pentamer. Molecular clusters with high conformational freedom like clusters of β-alanine can dissociate following different patterns. The stability of the clusters and some physico-chemical properties such as reactivity, can be studied directly in terms of dissociation energy. Therefore, for weakly bonded systems, the dissociation energies are significant quantities, giving important additional information about the stability. However, to compute the dissociation energies for all possible channels in systems with a large number of degrees of freedom, as neutral clusters of amino acids, is not a trivial task (see computational details).

Eq. (12.7) was used to calculate all possible scenarios of the dissociations. The different computed quantities and the importance of the corrections are also presented in the Tab. 12.1. For instance, the BSSE correction computed with the CP method is in the same energy range that the previously reported for hydrogen-bonded dimers (HF/HCN, HF/H_2O and HCCH/H_2O).[359] However, it has been shown that for weaker interaction energies the effect of BSSE a priori becomes larger, although the basis set used in this manuscript is designed to avoid this effect.[360]

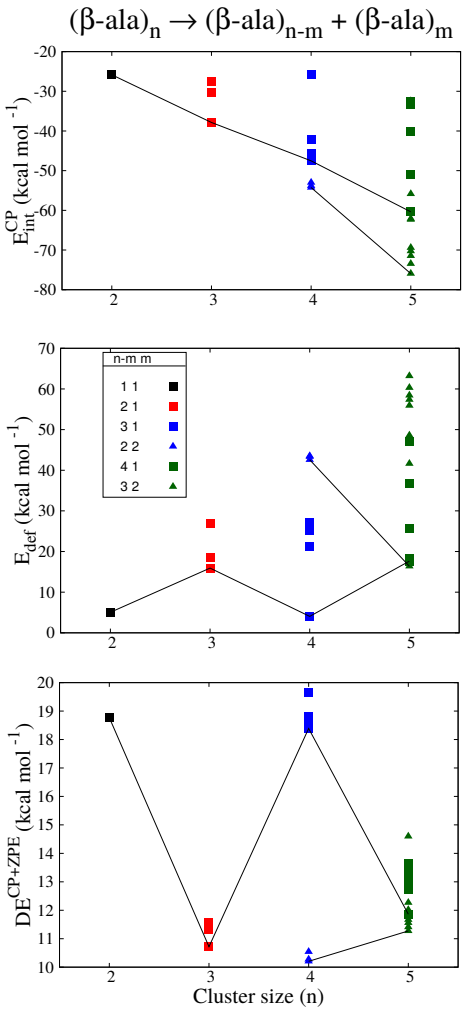


Figure 12.3: Interaction (E_{int}^{CP}), deformation (E_{def}) and dissociation (DE) energies in kcal·mol^{-1} as functions of the cluster size n, for the most stable neutral cluster $(\beta - ala)_n$ $n \leq 5$. The employed isomers for n, $n-m$ and m correspond to the most stable conformer (see Fig. 12.2). DE includes $BSSE$, and ZPE corrections.

Fig. 12.3 shows the variation of the interaction, deformation and dissociation energies as a function of the parent cluster size, n. The quantities were

computed removing one or two residues from the cluster:
$(\beta - ala)_n \rightarrow (\beta - ala)_{n-1} + \beta - ala$
$(\beta - ala)_n \rightarrow (\beta - ala)_{n-2} + (\beta - ala)_2$.
Also, it has to be mentioned that we considered all possibilities for each pattern i.e. 1 for dimer, 3 for trimer, 7 for tetramer and 15 for pentamer. The interaction energies become more negative as a function of n. This is due to the bigger sizes, where simply the probability of founding intermolecular interactions within the cluster is higher. We can also observe that the range of interaction energies becomes wider for bigger clusters, for pentamer it is found from ≈ -25 to -80 kcal mol^{-1}. Moreover, the E_{int}^{CP} computed values are stronger for the 2+2 and 2+3 channels than for the loss of 1 molecule.

Deformation energies are positive and fluctuate with the cluster size (see Fig. 12.3). The bigger values of E_{def} are observed for the (2+2) and (2+3) patterns, which means that more energy has to be given to the system to deform the geometry of the most stable dimer and trimer to the geometries shown in the tetramer and pentamer, respectively. In addition, Fig. 12.3 reveals that there is a systematic variation of deformation energy for the different dissociation channels.

Finally, the calculated DE^{CP+ZPE} is in all cases above 10 kcal mol^{-1} and shows a strong oscillatory behaviour with the cluster size (even-odd alternation). We can see that for a given cluster size the dissociation energies are larger, when the fragments do not suffer strong structural deformations during the dissociation, i.e., the dissociation energy increases when deformation energy decreases for a given channel. This shows that the structure of the fragments within the cluster is also important and strongly influences its stability.

Intermolecular hydrogen bonding

The intra and intermolecular non-covalent interactions for the most stable structures have been visualised using the NCI approach (see Fig.12.4). A closer look at the intermolecular interactions in the structures presented in Fig. 12.2 and Fig.12.4, indicates that they are HBs. Different atoms are involved in the intermolecular hydrogen bonding: O, N and C. We observe different types of HBs: amine-carboxyl (head to tail), carboxyl-carboxyl (tail to tail) and amine-amine (head to head) interactions. They can be classified using the hydrogen donor (D) atom and hydrogen acceptor (A) atoms denoted as $D-$H and H$---A$, respectively. The number of HBs, and its type, increases as the cluster size increases. The most stable clusters present head to tail and tail to head interactions: two kind of interconnections are observed according to the hydrogen donor-acceptor atoms (N-H$---$O or O-H$---$N), in which nitrogen or oxygen is given the hydrogen atom to the HB, respectively. Indeed, the most stable conformer of the dimer is a good illustration of cyclic structure with head to tail interactions, where we observe two HBs of the type O-H$---$N. Fig. 12.2 shows the presence of tail to tail hydrogen bonding (O-H$---$O) as well. For less stable dimers and also for bigger clusters it is a common that one carboxyl-carboxyl interaction is combined with other tail to

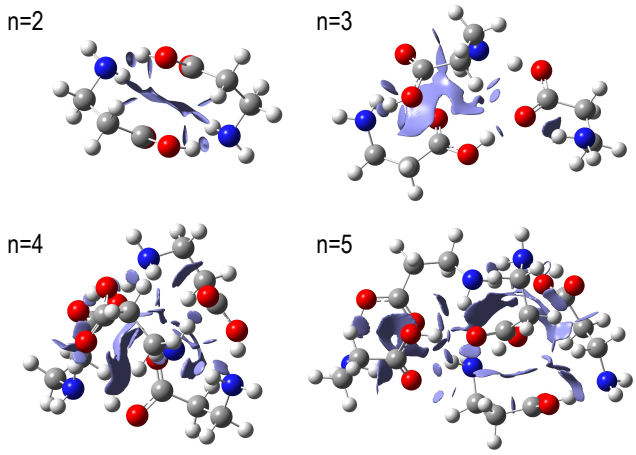


Figure 12.4: Reduced Density Gradient with an isovalue of 0.4 a.u. for the most stable structure of dimer, trimer, tetramer and pentamer. The non-covalent regions appear as lobes in the image. They have been computed with the Non-Covalent Interactions (NCI-plot3.0) visualisation tools[344;345]

tail HB, therefore lowering the relative energy of these structures (see appendices). Fig.12.2 shows a lower presence of amine-amine (head to head) interactions in the most stable isomers of each cluster size. Similarly, the presence of intermolecular amide-amide HBs has not been established in halogenated solvents at room temperature.[361]

The study of the intermolecular interactions, in particular the HBs presented in Fig. 12.2, together with an electron density analysis of these bonds allows us to understand the chemical stability of these systems. The topological analysis of the electron density at the bond critical point (BCP) is a very useful tool to characterise the nature of the HBs. The data gathered in the Fig. 12.2 suggests that the stability of the structures should increase with the cluster size, since the number of BCP increases. The number of ring critical points (RCP), also increases forming many-member ring structures except for the trimer. However, clusters composed of five molecules are not big enough to find more than two cage critical points (CCP). It is therefore observed that clusters prefer closed configurations bringing to mind cage type structures. The nature of intermolecular interactions was qualitatively studied by evaluating of the electron density (ρ) and the Laplacian of the electron density ($\nabla^2\rho$) in the BCP; this data is reported in Tab. 12.2.

Fig. 12.5 shows the relation between the intramolecular donor-hydrogen distance (r_{D-H}), electronic density (ρ) at the BCP, Laplacian of the electronic density ($\nabla^2\rho$) at the BCP and bond stabilisation energy (ϵ_{NBO}) as a function of the intermolecular hydrogen-acceptor distance (r_{H---A}). It is well established that the

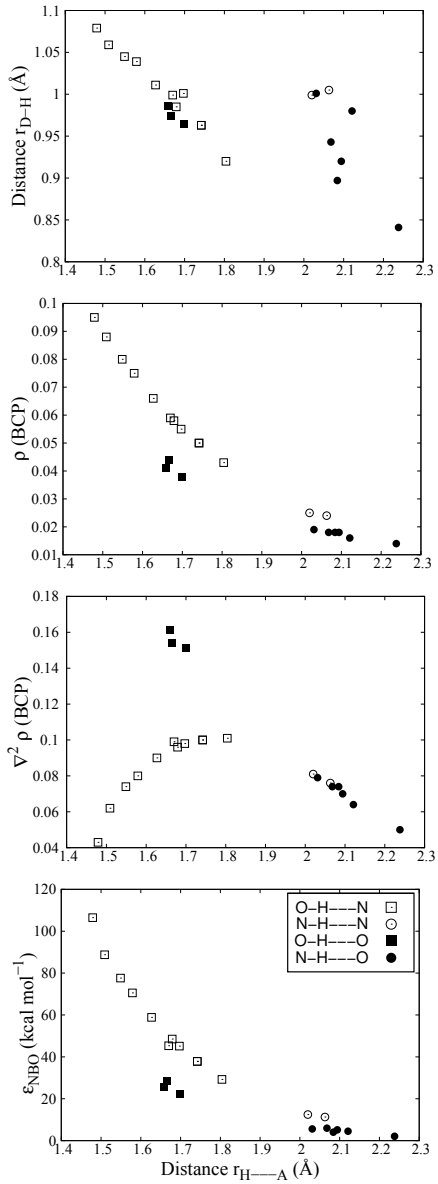


Figure 12.5: Intramolecular hydrogen donor-hydrogen distance (D−H), electron density (ρ), Laplacian of the electron density ($\nabla^2\rho$) at the bond critical point and bond stabilisation energy (ϵ_{NBO}) as functions of the intermolecular distance between hydrogen atom and the hydrogen acceptor atom ($\mathrm{r_{H---A}}$). Open squares and open circles correspond to HBs in which nitrogen is the acceptor atom. Solid squares and solid circles represent HBs in which oxygen is the acceptor atom. Squares correspond to the HBs in which oxygen is the donor atom. Circles represent the HBs in which nitrogen is the donor atom.

Table 12.2: Electron density (ρ) and Laplacian of the electron density ($\nabla^2\rho$) at the bond critical point (BCP) in a.u./e^3 and a.u./e^5, respectively; stabilisation energy (ϵ_{NBO}) (in kcal mol^{-1}) of each HB and total stabilisation energy ($\mathrm{E_{NBO}}$) for the most stable clusters of β-alanine. HBs indexes correspond to the numbers shown in Fig. 12.2a. The hydrogen donor atom (D) and the hydrogen acceptor atom (A) define the type of HB (D–H– – –A). Distance between hydrogen and hydrogen acceptor atom (r_{H---A}) is given in Å. With QT-AIM we did not find HB number 8 for pentamer.

Index	Type	ρ	$\nabla^2\rho$	r_{H---A}	ϵ_{NBO}	$\mathrm{E_{NBO}}$
		DIMER $(\beta - ala)_2$				
1	O-H– – –N	0.050	0.100	1.742	37.85	75.69
2	O-H– – –N	0.050	0.100	1.742	37.84	
		TRIMER $(\beta - ala)_3$				
1	O-H– – –N	0.088	0.062	1.509	88.76	
2	N-H– – –O	0.016	0.064	2.121	4.50	144.98
3	O-H– – –N	0.043	0.101	1.804	29.23	
4	O-H– – –O	0.038	0.151	1.699	22.49	
		TETRAMER $(\beta - ala)_4$				
1	O-H– – –N	0.080	0.074	1.549	77.61	
2	O-H– – –O	0.044	0.154	1.666	28.33	
3	N-H– – –O	0.018	0.074	2.068	5.96	
4	N-H– – –O	0.014	0.050	2.238	2.06	259.86
5	N-H– – –O	0.018	0.074	2.084	4.03	
6	N-H– – –N	0.025	0.081	2.020	12.47	
7	O-H– – –N	0.066	0.090	1.627	58.86	
8	O-H– – –N	0.075	0.080	1.579	70.54	
		PENTAMER $(\beta - ala)_5$				
1	O-H– – –N	0.058	0.096	1.679	48.51	
2	N-H– – –N	0.024	0.076	2.063	11.28	
3	N-H– – –O	0.019	0.079	2.031	5.59	
4	O-H– – –O	0.041	0.161	1.659	25.63	
5	O-H– – –N	0.055	0.098	1.697	45.14	295.16
6	N-H– – –O	0.018	0.070	2.094	5.15	
7	O-H– – –N	0.095	0.043	1.479	106.48	
8	O-H– – –C	x	x	2.418	2.06	
9	O-H– – –N	0.059	0.099	1.670	45.32	

distance varies with the kind and nature of the HB.[362–365] The proton affinity of the nitrogen and oxygen atoms (mainly O-H$---$N and N-H$---$O) stabilises the clusters in a widely spread range of intermolecular HBs distances. The distances of different types of HB within all clusters sizes range from 1.48 to 1.80 Å, from 1.68 to 1.70 Å, from 2.02 to 2.06 Å and from 2.03 to 2.24 Å for O-H$---$N, O-H$---$O, N-H$---$N and N-H$---$O interactions, respectively. One can notice that for the shortest intermolecular distances, $r_{H---A} \leq 1.9$ Å, the oxygen is the donor atom to the HB. For the oxygen donor atom, in contrast to nitrogen, a direct coupling between the intra and intermolecular distances is also observed i.e. as bigger the r_{H---A} as smaller the r_{D-H} distance and vice versa.

Analysing the electron density in the BCP as a function of the intermolecular distance of the corresponding HB we observe two main patterns. Higher electron densities in the BCP are observed for oxygen donor atom and accordingly, shorter r_{H---A} distances were shown. For nitrogen as a donor atom this trend is different: ρ remains almost constant as a function of r_{H---A}. In general, the electron density decreases exponentially with r_{H---A}, for both nitrogen and oxygen as hydrogen acceptors. Therefore, the presence of different types of sequence-specific HB, in a wide range of intermolecular distances, contribute differently to the stability of clusters of β-alanine, as it is also shown by the electron density at the BCP.

The value of the Laplacian of the electronic density at the corresponding BCP increases only for O-H$---$N type of HB. For other types of interactions $\nabla^2\rho$ decreases as a function of r_{H---A} distance.

In general, these findings are consistent with a molecular orbital description: the more electronegative hydrogen donor atom, the longer r_{D-H} distance, and therefore a smaller r_{H---A} and stronger stabilisation effect is observed in ρ and $\nabla^2\rho$ at the BCP.

The stabilisation energies for the most stable isomer for each cluster size are presented in Tab. 12.2. The main contribution to the stabilisation energy comes from lone pair (LP) orbitals. The type of interactions mainly observed between residues are: (i) nitrogen-LP with $\sigma_{OH}{}^*$ orbitals, (ii) oxygen-LP with $\sigma_{OH}{}^*$ orbitals and (iii) nitrogen-LP with $\sigma_{NH}{}^*$ orbitals (see appendices for more details). These interactions are good textbook examples of hyperconjugation, commonly known as stabilising effect (non-diagonal elements of the Fock matrix in the NBO basis) between unoccupied and occupied orbitals of a Lewis-type structure (LP $\rightarrow \sigma^*$).[154]

Similarly to the electronic densities, the stabilisation energies decrease exponentially with the hydrogen-acceptor distance (r_{H---A}). The HB stabilisation energy depends on the different donor acceptor interaction i.e. it increases in the sequence: N-H$---$O $<$ N-H$---$N $<$ O-H$---$O $<<$ O-H$---$N. The strongest HB stabilisation energy is found for pentamer (HB labeled 7 in Fig. 12.2). The main stabilisation effects of this interaction are $LP_N \rightarrow \sigma_{OH}{}^*$ (93.72 kcal mol^{-1}) and $\sigma_{NH} \rightarrow \sigma_{OH}{}^*$ (4.48 kcal mol^{-1}). Such strong stability is reflected in a very short intermolecular distance, $r_{H---A} = 1.479$ Å.

The geometry of the most stable dimer is oriented into a ring structure. The interactions in this structure stabilise the cluster with a total energy of E_{NBO}=75.7

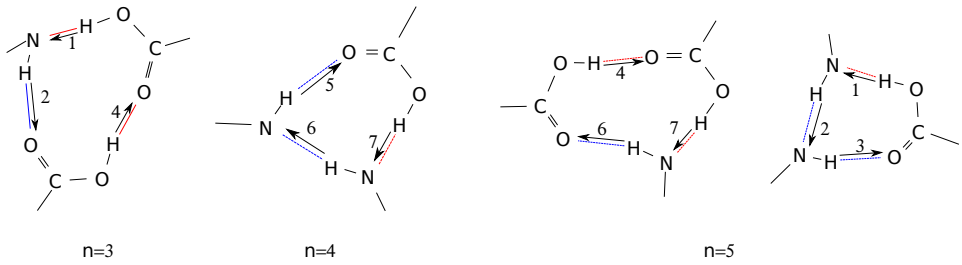


Figure 12.6: Scheme of the clockwise or anticlockwise interactions for trimer, tetramer and pentamer. The directions of the HB from the donor to acceptor is shown with an arrow.

kcal mol^{-1} (two O-H$-$ $-$ $-$N bonds). Interaction energies have been previously studied for L-alanine dimer, where small hydrogen-bonded rings were also identified.[366]

The role of H atoms and HBs in ring structures has been reported previously.[367;368] The rings found in the most stable clusters with HBs directly involved in a cyclic structure are presented in Fig. 12.6. To simplify this analysis, we ignore those structures separated by more than two covalent bonds, ending up in the planes connected by 3 HBs. The four considered hydrogen-bonded cyclic structures follow the same hydrogen bonding arrangements, i.e. the so called "clockwise" or "anticlockwise" donation $D-$H$-A$ pattern.[369] Quasi-pentagonal hydrogen-bonded rings are observed for tetramer and pentamer. Three residues of the trimer and the pentamer create quasi-hexagonal hydrogen-bonded structures,[370] with an analogous geometry to a twist-boat conformation of cyclohexane. The extra stability of the pentamer is governed by the presence of both rings forming an analogue

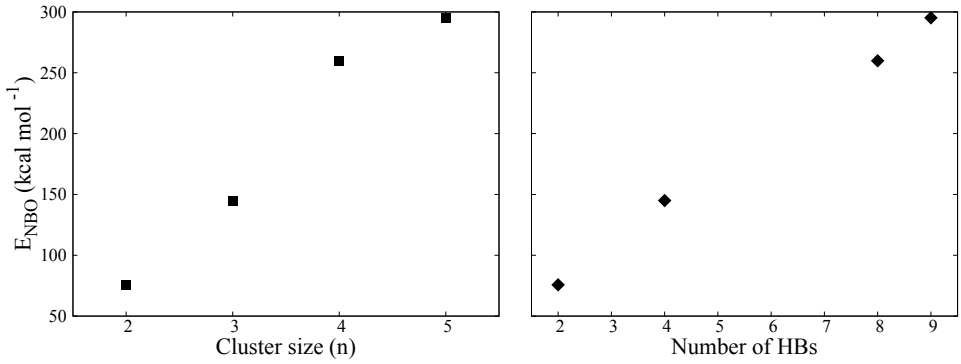


Figure 12.7: Stabilisation energy (E_{NBO}) as a function of the cluster size n and as a function of the number of HBs in the structure, for the most stable dimer, trimer, tetramer and pentamer.

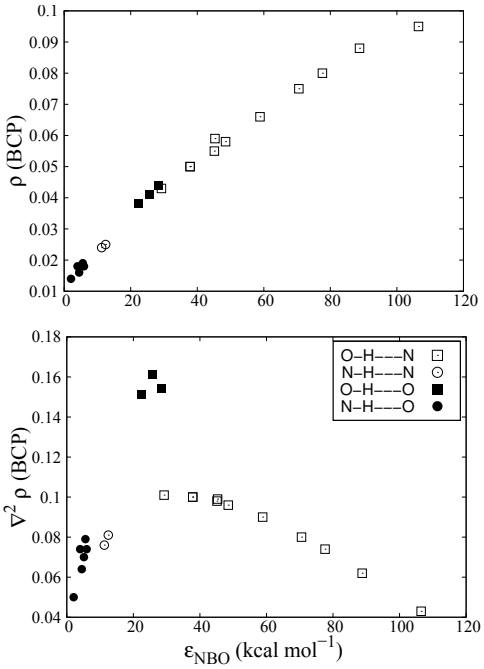


Figure 12.8: Electron density and Laplacian of the electron density at the bond critical point as functions of the stabilisation energy (E_{NBO}). Open squares and circles correspond to HB in which nitrogen is the acceptor atom, solid squares and circles - oxygen is the acceptor atom, squares - oxygen is the donor atom, circles - nitrogen is the donor atom.

structure to inner cavities found inside proteins, which strongly influence their folding.[371]

The direction of the HB in the rings could be an indication of cooperative effects. Cooperative effects were observed previously in halogenated weakly bonded[372;373] and other non-covalently interacting systems[374–376] but never for pure clusters of amino acids. Cooperativity of amidic HBs (N-H− − −O=C, in this manuscript defined as N-H− − −O) has been previously observed experimentally[377–380] and theoretically.[381;382] It has been shown that the presence of this type of HB contributes enormously the stability of α-helices.[383] Recently, α-helical cooperativity has been also demonstrated in an intact protein.[384] The presence of amidic HBs and the ring structures (see Fig. 12.6) supports the idea of the cooperative effects in the clusters of β-amino alanine molecules. To confirm this hypothesis we calculated the ϵ_{NBO} of HBs involved in the ring structures without the influence of the other HBs (i.e. removing surrounding residues), finally ending up with only two residues for each case, and keeping the optimised geometry of the 12 possibilities considered. It was observed that the presence of amidic HBs reinforce other HB interactions as was reported previously for α-helices. The most meaningful example of the HB

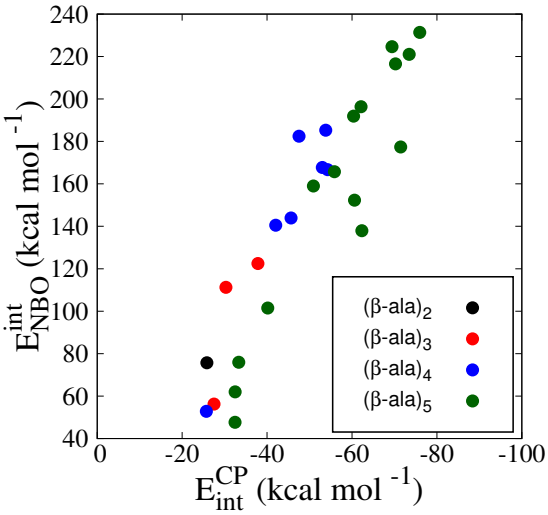


Figure 12.9: Stabilisation energy involved in the dissociation (E^{int}_{NBO}) as a function of the corresponding interaction energy (E^{CP}_{int}).

strengthening is found for the 6-member ring in pentamer, where the cooperativity of amidic and O− − −O-H bonds significantly increases ϵ_{NBO} of the 7^{th} HB by 12.4% (11.8 kcal mol^{-1}). In general all the HBs energies (ϵ_{NBO}) in the analysed ring structures present cooperativity with stabilisation up to 22.6% higher than for non interacting residues (see appendices for more details). Therefore, the role of the N-H− − −O HB in stabilising the clusters by cooperative effects is crucial. Similar synergy of different types of HBs in polyglycine models of antiparallel β-sheets has been found to strengthen or weaken a particular intermolecular HB.[385] Intramolecular HB in protonated disubstituted ethenes, benzenes and naphthalenes have also shown similar relationships.[386] Moreover, a key aspect for prediction of the stability and a proper description of weakly bonded systems is not only the number and the type of the HBs (their independent stabilisation energies, ϵ_{NBO}) but also their complex relative dependencies enclosed in the HBs rings. All that become the clear manifestation of cooperative effects.[387;388]

Fig. 12.7 shows that the total stabilisation energy increases linearly with the number of HBs and, consequently, with the cluster size. A similar linear dependence is observed in the Fig. 12.8: the larger the electronic density in the BCP the stronger stabilisation interactions. In the same figure we see that the Laplacian of the electronic density increases for stabilisation energies up to 30 kcal mol^{-1} and decreases almost linearly for higher values of ϵ_{NBO}. Therefore, we observe direct relations between ρ and $\nabla^2\rho$ (calculated with QT-AIM) and ϵ_{NBO} (calculated with NBO). This also demonstrates that an enhancement of the local electronic density at the BCP and a positive Laplacian of the electronic density determine

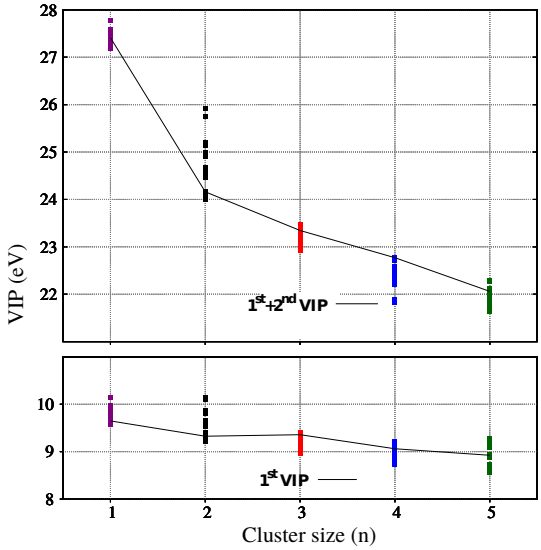


Figure 12.10: First and first+second vertical ionisation potentials (VIPs) for all studied isomers of $(\beta - ala)_n$ $n \leq 5$. Dashed and solid lines correspond to the 1^{st} and 1^{st}+2^{nd} VIPs, respectively, of the most stable neutral conformer of each cluster size n.

the stabilisation energy of the different types of HB and reveal their role during the cluster formation.

Fig. 12.9 shows the relation between the interaction energy (E_{int}^{CP}) and the corresponding NBO stabilisation energy (E_{NBO}^{int}), computed as the sum over all the HB stabilisation energies involved in a certain dissociation pattern. In this way, two different quantities from two different approaches, E_{NBO}^{int} and E_{int}^{CP}, are compared. We observe an almost linear behaviour showing the strong relation between them. Notice that the local description of bonds obtained with the NBO method overestimate the stabilisation effects.

Vertical ionisation potentials

Results in Fig. 12.10 address the issue of the cluster stability against ionisation. For each cluster size we calculated the 1^{st} and 1^{st}+2^{nd} vertical ionisation potentials (VIP) i.e. the energy required to remove one or two electrons keeping the geometry of the neutral cluster. Obviously, removal of one electron requires significantly less energy than removal of two electrons. We can notice as a general trend that the VIP decreases with the size of the system. The bigger the system, the easier it is to accommodate the excess of positive charge. First VIP for pentamer appears around 1 eV lower than for the monomer, which means that ionisation of bigger clusters demands less energy. A big increase is observed for the 1^{st}+2^{nd} VIP between the

monomer and the dimer, a double ionisation of the dimer requires $\approx$ 3 eV less energy than the for monomer. For bigger clusters 1^{st}+2^{nd} VIP decreases linearly with the size. Furthermore, the different isomers of each cluster size present very wide range of ionisation energy thresholds. The most stable isomer of the monomer and dimer shows the lowest 1^{st} VIP among all monomers and dimers, for the rest of the cases this trend is not observed.

12.4 Conclusions

We have presented a theoretical study of neutral clusters of $\beta-$alanine molecules $(\beta - ala)_n$ in the gas phase with sizes $n \leq 5$. We have first introduced a computational strategy to deal with the large number of possible isomers that these systems may adopt. It consists in performing classical molecular dynamics simulations using the GAFF force field with different temperatures and, after a first screening, we optimise the selected structures with density functional theory. We have also presented a benchmarking study for a proper selection of the density functional theory method that best reproduce these systems. Based on energetic, geometric and bond-properties criteria, the M06-2X functional in combination with the 6-311++G(d,p) basis was chosen as the best level of theory to describe clusters of amino acids in the gas phase. With this level of theory we have evaluated the stability of the clusters in terms of the intermolecular interactions between the residues. Strong HBs are responsible of such stability. In particular cooperative effects in ring structures strengthen these bonds. An in-deep study of wave function properties and interaction energies within these bonds reveal that the nature of the donor and acceptor atoms (D-H—A) determines the bond strength. We have also computed dissociation energies including zero point energy and basis set superposition corrections: the dissociation energy shows an oscillatory (even-odd) behaviour with the cluster size. First and second vertical ionisation potentials have been evaluated as well; both decrease with the cluster size.

Part IV

Conclusions and perspectives

Chapter 13

Conclusions and perspectives

Conclusions

In this thesis we have studied theoretically the behaviour and evolution of thymidine, glycine, β-alanine, γ-aminobutyric acid and N-acetylglycine after an ionising collision with a highly charged atomic ion in the gas phase. This thesis is also the outcome of the close collaboration with the group that perform such experiments. We have investigated the stability, fragmentation dynamics and different de-excitation pathways characterising the decay of the positively charged molecules. Our theoretical predictions are in good agreement with the experimental measurements. This thesis also includes a pure theoretical study of neutral clusters of β-alanine molecules, focusing on their structure and stability in terms of weak interactions. The following conclusion are derived:

- The energy transferred in the collision, for a single electron capture was determined qualitatively and quantitatively for thymidine and glycine, respectively.
  - In the case of singly-charged thymidine, the experimental measurements have shown an excitation energy distribution with the maximum of $\approx$2.5 eV and a broad distribution of $\approx$ 6 eV. Thereafter, such an excitation energy distribution have been used in the theoretical modelling for the rest of the systems studied in this thesis. The theoretical calculations shows that (i) an energy deposit of [1-5] eV explain the observed peaks in the mass spectra, (ii) the charge localization is correlated with the molecular orbital involved in the ionisation.
  - For singly-charged glycine it has been qualitatively showed that the charge state of different projectiles determine the energy transferred to the molecule according to the calculated pathways on the potential energy surface and the observed peaks in the experiments. Differences in the energy-transfer is clearly reflected in the neutral water emission. The peak corresponding to the [glycine without H_2O]$^+$ is stronger when the charge state of the projectile increases. It suggests that the cap-

ture is more gentle and the excitation energy transferred to the target is smaller for a highly charged ion.

- Different de-excitation pathways characterise the decay of the doubly charged lineal amino acids: glycine, β-alanine and γ-aminobutyric acid after interaction with highly charged ions. In all of the cases the most significant fragmentation channels corresponds to molecular breaking into two singly charged fragments, the so-called Coulomb explosion. This process leads to:
  - $C_{carboxyl} - C_\alpha$ bond cleavage for glycine,
  - $C_{carboxyl} - C_\alpha$ or $C_\alpha - C_\beta$ bond breaking for β-alanine,
  - two body $C_\beta - C_\gamma$ or three body $C_{carboxyl} - C_\alpha + C_\beta - C_\gamma$ fragmentation for GABA.
- In competition with the Coulomb explosion, ultrafast H migration, OH migration and cyclisation processes take place, leading to unexpected fragmentation channels. The formation of stable dications followed by the evaporation of neutral moieties has been observed:
  - Doubly charged fragments have been found due to the fast (a few $\approx$10 of fs) hydrogen migration for glycine.
  - In addition to H migration, OH migration has been also detected ($\approx$100 fs) for β-alanine.
  - In addition to H migration, slower processes ($\approx$100-200 fs) such as many-member ring formations have been found for GABA. Due to these reactions stable doubly charged oxygen and nitrogen reactive species have been formed.
- For singly charged N-acetylglycine it has been shown that the peptide bond provides additional stability against the fragmentation. The tautomerisation processes have been identified due to the presence of this bond. They produce the cationic diol structures, with high stability, which are responsible for the long-lived metastable species observed in the experiments. Such processes may occur in other systems containing a peptide bond and has to be considered to obtain a complete picture of the fragmentation dynamics in peptides and proteins.
- For neutral cluster of β-alanine molecules a new computational strategy is proposed. It has been shown that complete and accurate conformational searches for weakly bonded clusters can be performed using classical molecular dynamics simulations with general amber force field if one perform careful parameters check. Moreover, it has been concluded that the M06-2X functional is the most accurate DFT choice to calculate properties of these systems. Strong cooperative effects have been found due to presence of different types of hydrogen bonds. They present more likely intramolecular

stabilisation leading to cage type structures, rather than evolve into linear weakly bonded chains.

Chapter 14

Conclusions and perspectives
Conclusiones

En esta tesis hemos estudiado mediante métodos teóricos el comportamiento y la evolución de la timidina, glicina, β-alanina, ácido γ-aminobutírico y N-acetilglicina después de una colisión ionizante con un ion atómico altamente cargado en fase gas. Esta tesis es también el resultado de una estrecha colaboración con el grupo que realiza tales experimentos. Hemos investigado la estabilidad, la dinámica de fragmentación y diferentes vías de desexcitación que caracterizan la desintegración de dichas moléculas cargadas positivamente. Nuestras predicciones teóricas están en buen acuerdo con las medidas experimentales. Esta tesis también incluye un estudio teórico puro de agregados moleculares neutros de β-alanina, centrándose en su estructura y estabilidad en términos de interacciones débiles. Se derivan de este trabajo las siguientes conclusiones:

- La energía transferida en la colisión, para la captura de un solo electrón se determinó cualitativa y cuantitativamente para timidina y glicina, respectivamente.
  - En el caso de timidina simplemente cargada, las medidas experimentales han mostrado una distribución de energía de excitación con un máximo de ≈ 2.5 eV y una distribución ancha de ≈ 6 eV. Posteriormente, dicha distribución de energía de excitación se ha utilizado en las simulaciones teóricas para el resto de los sistemas estudiados en esta tesis. Los cálculos teóricos demuestran que (i) un depósito de energía de [1-5] eV explica los picos observados en los espectros de masas, (ii) la localización de carga se correlaciona con los diferentes orbitales moleculares implicados en la ionización.
  - Para la glicina simplemente cargada, se ha demostrado cualitativamente que el estado de carga de diferentes proyectiles determina la energía transferida a la molécula de acuerdo con las vías calculadas sobre la superficie de energía potencial y los picos observados en los experimentos. Las diferencias en la transferencia de energía se reflejan claramente en

la emisión de una molécula de agua neutra. El pico correspondiente a la [glicina sin H_2O]$^+$ es mayor cuando el estado de carga del proyectil aumenta. Esto sugiere que la captura es más suave y la energía de excitación transferida a la molécula es menor para un ion altamente cargado.

- Diferentes vías de desexcitación caracterizan la descomposición de los aminoácidos lineales doblemente cargados: glicina, β-alanina y ácido γ-aminobutírico después de la interacción con iones altamente cargados. En todos los casos, los canales de fragmentación más significativos corresponden a la ruptura molecular en dos fragmentos cargados individualmente, la llamada explosión de Coulomb. Este proceso conduce a:
  - ruptura del enlace $C_{carboxil} - C_\alpha$ para glicina,
  - ruptura de los enlaces $C_{carboxyl} - C_\alpha$ or $C_\alpha - C_\beta$ para β-alanina,
  - dos rupturas en dos fragmentos $C_\beta - C_\gamma$ o en tres fragmentos $C_{carboxyl} - C_\alpha + C_\beta - C_\gamma$ para GABA.
- En competencia con la explosión de Coulomb, se producen migraciones ultrarápidas de H, migración de OH y procesos de ciclización, lo que lleva a canales de fragmentación inesperados. Se ha observado la formación de dicationes estables como consecuencia de la evaporación de fragmentos neutros después de la reorganización molecular:
  - Se han encontrado fragmentos doblemente cargados debido a la migración rápida de hidrógeno ($\approx$ 10 fs) para la glicina.
  - Además de la migración H, también se ha detectado la migración de OH ($\approx$ 100 fs) para β-alanina.
  - Además de la migración de H, se han encontrado procesos más lentos ($\approx$ 100-200 fs) como formaciones de anillos de distintos tamaños para GABA. Debido a estas reacciones, se han formado especies reactivas estables doblemente cargadas de oxígeno y de nitrógeno.
- Para la N-acetilglicina simplemente cargada se ha demostrado que el enlace peptídico proporciona estabilidad adicional frente a la fragmentación. Se han identificado procesos de tautomerización debido a la presencia de este enlace. Se han encontrado estructuras de dioles catiónicos, con alta estabilidad, que son responsables de las especies metaestables de larga vida observadas en los experimentos. Estos procesos pueden ocurrir en otros sistemas que contienen un enlace peptídico y tienen que ser considerados para obtener una imagen completa de la dinámica de fragmentación en péptidos y proteínas.
- Para agregados neutros de β-alanina, se propone una nueva estrategia computacional. Se ha demostrado que búsquedas conformacionales completas

y precisas para los agregados débilmente enlazados pueden llevarse a cabo usando simulaciones dinámica molecular clásica, con el potencial AMBER si uno realiza una verificación cuidadosa de los parámetros empleados. Por otra parte, se ha concluido que el funcional M06-2X es la elección de DFT más precisa para calcular propiedades de estos sistemas. Se han encontrado fuertes efectos cooperativos debido a la presencia de diferentes tipos de enlaces de hidrógeno. Estos sistemas presentan una estabilización intramolecular que conduce a estructuras tipo jaula, en lugar de evolucionar en cadenas lineales débilmente enlazadas.

Chapter 15

Conclusions and perspectives
Perspectives

In the future, we are planning to extensively study the importance of the electronic excited states in the fragmentation of ionised biomolecules. Time-dependent DFT and more accurate techniques, such as CASSCF with surface hopping molecular dynamics method will be applied, to get deeper insight into charge migrations during the fragmentation processes. We will also investigate the mechanisms behind the peptide bond formation in the ionised clusters of β-alanine molecules. Future studies will also aim to the investigation of unimolecular decomposition of more complex systems like solvated clusters of β-alanine and/or Br-uracil.

Molecular dynamics in the electronic excited states

Different types of experiments based on ionisation with light sources, such as synchrotron radiation ultrashort or laser pulses suggest that particular fragmentation channels can not be explained without considering of electronic excited states. [131;170] Therefore, in addition to the benchmarking study performed for the fragmentation of the S-α-alanine molecule in the electronic ground state, we performed also a detailed statistical analysis including the first excited state. To this, we used surface hopping non-adiabatic linear-response time dependent density functional theory (SH-LR-TDDFT). [42;132;389] The preliminary results of this study are presented below. The comparison of BOMD and SH-LR-TDDFT disentangle the importance of the excited states for particular exit channels. The general picture of fragmentation reproduce the previous experimental study using highly-charged-ion-induced fragmentation. [211] However, long timescale *ab initio* molecular dynamics (AIMD) simulations explain the mechanism involving double hydrogen migration process leading to NH_4^++$COOH^+$. The importance of the fragmentation channels including electronic excited states explains recent experimental measurements performed in X-ray induced fragmentation experiments (private communication by Prof. Edwin Kukk).

Additional multireference theoretical studies based on molecular dynamics sim-

ulations with surface hopping method[133] have been also performed in order to provide interpretation of recent experimental measurements of ion induced fragmentation of betaine (private communication by Dr. Patrick Rousseau et al.). Our simulations indicate that including the first 3 excited states is necessary to reproduce the experimental measurements. These preliminary studies show strong couplings between the electronic excited states (ground/first, first/second and second/third), which drive different charge transfer fragmentation dynamics of this system. Such intricate correlation is responsible for the transfer of excitation energy towards emission of CO_2. Partially delocalised electronic charge, already favours energy transfer to this part of the molecule in the first femtoseconds of the simulations. On the other hand, ionising in the ground state does not show this trend. Emerging experimental data indicate that the fragmentation in the ground state is not fully described by a multiconfigurational ground state wave function picture.

Positively-charged molecular clusters

The theoretical study of ionised clusters of amino acids is of particular interest since it offers an excellent opportunity to unravel structural and electronic properties of biological material as a function of the cluster size including different charge states. Ionisation may induce the formation of covalent bonds between residues or the dissociation and fragmentation of the cluster. The key factor to understand the competition between these chemical processes lay in the heart of the weak interactions between residues. For instance it is well known that on atomic clusters, at low interaction energies only single atoms can be ejected. If the interaction energies increase, then dimers or trimers can easily dissociate. However, on molecular clusters, stabilised with H bonds, the problem is more complex. Does the size-dependent effects appear in the peptide bond formation for ionised $(\beta\text{-ala})_n^{q+}$ clusters? One can analyse the experimental coincidences with the peaks associated with the peptide bond formation and use this information to explore the corresponding regions of the PES. Combining the calculations of adequate pathways on the PES with experimental evidences may be a general strategy to understand and explain new covalently bonded chemical species. Molecular dynamics in the first steps can provide a complementary picture of the PES exploration.

In recent experiments of (private communication by Dr. Patrick Rousseau et al.) after ionisation of β-alanine clusters in the gas phase with high energy O^{3+} ions, emission of neutral moieties of mass 18 amu is observed. Considering that the mass corresponds to the removal of the neutral water, we expect to observe the formation of a peptide bond. Creating a peptide bond by removing a water molecule in a cluster of amino acids could be associated with hydrophobic and hydrophilic properties of their side chain. Before ionisation, the neutral clusters of amino acids are weakly bonded forming cage structures as shown in this thesis. After the ionisation, charged and polar side chains in the cluster become hydrophilic

and can loss water and also "drain" out the water molecule from the core of the cluster. Because of that, the geometry of the clusters plays a dominant role in the formation of the peptide bond. The formation of covalent bonds like C-C can also take place leading to functionalising derivatives.[390] We plan to run MD and PES simulations to understand ion induced reactivity inside the clusters.

Bibliography

Bibliography
BIBLIOGRAPHY

[1] A. Stern and N. H. Lindner. Topological quantum computation-from basic concepts to first experiments. *Science*, 339, 2013.

[2] A. Szabo and N. S. Ostlund. *Modern Quantum Chemistry: Intro to Advanced Electronic Structure Theory*. Dover Publications, 1996.

[3] H. da Silva Jr, J. Oller, M. Gatchell, M. H. Stockett, P.-A. Hervieux, L. Adoui, M. Alcami, B. A. Huber, F. Martin, H. Cederquist, et al. Multiple electron capture, excitation, and fragmentation in C^{6+}- C_{60} collisions. *Physical Review A*, 90:032701, 2014.

[4] L. Sanche. Low energy electron-driven damage in biomolecules. *The European Physical Journal D, - Atomic, Molecular, Optical and Plasma Physics*, 35:367–390, 2005.

[5] D. Schardt, T. Elsässer, and D. Schulz-Ertner. Heavy-ion tumor therapy: Physical and radiobiological benefits. *Reviews of Modern Physics*, 82:353–425, 2010.

[6] P. López-Tarifa, M.-A. Hervé du Penhoat, R. Vuilleumier, M.-P. Gaigeot, I. Tavernelli, A. Le Padellec, J.-P. Champeaux, M. Alcamí, P. Moretto-Capelle, F. Martín, and M.-F. Politis. Ultrafast nonadiabatic fragmentation dynamics of doubly charged uracil in a gas phase. *Physical Review Letters*, 107:023202, 2011.

[7] P. López-Tarifa, M.-P. Gaigeot, R. Vuilleumier, I. Tavernelli, M. Alcamí, F. Martín, M.-A. Hervé du Penhoat, and M.-F. Politis. Ultrafast damage following radiation-induced oxidation of uracil in aqueous solution. *Angewandte Chemie International Edition*, 52:3160–3163, 2013.

[8] E. R. Stadtman. Protein oxidation and aging. *Science*, 257:1220–1224, 1992.

[9] D. A. Butterfield and D. Boyd-Kimball. The critical role of methionine 35 in Alzheimer's amyloid β-peptide induced oxidative stress and neurotoxicity. *Biochimica et Biophysica Acta. Proteins and Proteomics*, 1703:149–156, 2005.

[10] P. Ehrenfreund, D. P. Glavin, O. Botta, G. Cooper, and J. L. Bada. Extraterrestrial amino acids in orgueil and ivuna: Tracing the parent body of CI type carbonaceous chondrites. *Proceedings of the National Academy of Sciences*, 98:2138–2141, 2001.

[11] H. R. Pfaendler, J. Gosteli, R. B. Woodward, and G. Rihs. Structure, reactivity, and biological activity of strained bicyclic β-lactams. *Journal of the American Chemical Society*, 103:4526–4531, 1981.

[12] Alcaide B. and Almendros P. β-lactams as versatile synthetic intermediates for the preparation of heterocycles of biological interest. *Current Medicinal Chemistry*, 11:1921–1949, 2004.

[13] B. Fayard, A. Touati, F. Abel, M. A. H. du Penhoat, I. Despiney-Bailly, F. Gobert, M. Ricoul, A. L'Hoir, M. F. Politis, M. A. Hill, D. L. Stevens, L. Sabatier, E. Sage, D. T. Goodhead, and A. Chetioui. Cell inactivation and double-strand breaks: The role of core ionizations, as probed by ultrasoft x rays. *Radiation Research*, 157:128–140, 2002.

[14] B. Gervais, M. Beuve, G. Olivera, and M. Galassi. Numerical simulation of multiple ionization and high LET effects in liquid water radiolysis. *Radiation Physics and Chemistry*, 75:493 – 513, 2006.

[15] E. R. Williams. Proton transfer reactivity of large multiply charged ions. *Journal of Mass Spectrometry*, 31:831–842, 1996.

[16] T.-Y. Huang and S. A. McLuckey. Gas-phase chemistry of multiply charged bioions in analytical mass spectrometry. *Annual Review of Analytical Chemistry*, 3:365–385, 2010.

[17] S. A. McLuckey and J. L. Stephenson. Ion/ion chemistry of high-mass multiply charged ions. *Mass Spectrometry Reviews*, 17:369–407, 1998.

[18] D. K. Bohme. Multiply-charged ions and interstellar chemistry. *Physical Chemistry Chemical Physics*, 13:18253–18263, 2011.

[19] D. Schröder and H. Schwarz. Generation, stability, and reactivity of small, multiply charged ions in the gas phase. *The Journal of Physical Chemistry A*, 103:7385–7394, 1999.

[20] K. Lammertsma, P. von Ragué Schleyer, and H. Schwarz. Organic dications: Gas phase experiments and theory in concert. *Angewandte Chemie International Edition*, 28:1321–1341, 1989.

[21] D. Schröder. Coulomb explosions and stability of multiply charged ions in the gas phase. *Angewandte Chemie International Edition*, 43:1329–1331, 2004.

[22] J. Roithová and D. Schröeder. Bond-forming reactions of molecular dications as a new route to polyaromatic hydrocarbons. *Journal of the American Chemical Society*, 128:4208–4209, 2006.

[23] C. L. Ricketts, D. Schröder, C. Alcaraz, and J. Roithová. Growth of larger hydrocarbons in the ionosphere of Titan. *European Journal of Chemistry*, 14:4779–4783, 2008.

[24] U. J. Meierhenrich, G. M. M. n. Caro, J. H. Bredehöft, E. K. Jessberger, and W. H.-P. Thiemann. Identification of diamino acids in the Murchison meteorite. *Proceedings of the National Academy of Sciences of the United States of America*, 101:9182–9186, 2004.

[25] C. Chyba and C. Sagan. Endogenous production, exogenous delivery and impact-shock synthesis of organic molecules: an inventory for the origins of life. *Nature*, 355:125–132, 1992.

[26] J. P. Dworkin, D. W. Deamer, S. A. Sandford, and L. J. Allamandola. Self-assembling amphiphilic molecules: Synthesis in simulated interstellary precometary ices. *Proceedings of the National Academy of Sciences*, 98:815–819, 2001.

[27] M. P. Bernstein, J. P. Dworkin, S. A. Sandford, G. W. Cooper, and L. J. Allamandola. Racemic amino acids from the ultraviolet photolysis of interstellar ice analogues. *Nature*, 416:401–403, 2002.

[28] G. M. Muñoz Caro, U. J. Meierhenrich, W. A. Schutte, B. Barbier, A. Arcones Segovia, H. Rosenbauer, W. H.-P. Thiemann, A. Brack, and J. M. Greenberg. Amino acids from ultraviolet irradiation of interstellar ice analogues. *Nature*, 416:403–406, 2002.

[29] V. Blagojevic, S. Petrie, and D. K. Bohme. Gas-phase syntheses for interstellar carboxylic and amino acids. *Monthly Notices of the Royal Astronomical Society*, 339:L7–L11, 2003.

[30] Y.-J. Kuan, S. B. Charnley, H.-C. Huang, W.-L. Tseng, and Z. Kisiel. Interstellar glycine. *The Astrophysical Journal*, 593:848, 2003.

[31] L. E. Snyder, F. J. Lovas, J. M. Hollis, D. N. Friedel, P. R. Jewell, A. Remijan, V. V. Ilyushin, E. A. Alekseev, and S. F. Dyubko. A rigorous attempt to verify interstellar glycine. *The Astrophysical Journal*, 619:914, 2005.

[32] C. Barrientos, P. Redondo, L. Largo, V. M. Rayón, and A. Largo. Gas-phase synthesis of precursors of interstellar glycine: A computational study of the reactions of acetic acid with hydroxylamine and its ionized and protonated derivatives. *The Astrophysical Journal*, 748:99, 2012.

[33] A. Rimola, M. Sodupe, and P. Ugliengo. Computational study of interstellar glycine formation occurring at radical surfaces of water-ice dust particles. *The Astrophysical Journal*, 754:24, 2012.

[34] S. Maeda and K. Ohno. Generation mechanisms of amino acids in interstellar space via reactions between closed-shell species: Significance of higher energy isomers in molecular evolution. *The Astrophysical Journal*, 640:823, 2006.

[35] P. McCaldon and P. Argos. Oligopeptide biases in protein sequences and their use in predicting protein coding regions in nucleotide sequences. *Proteins*, 4:99–122, 1988.

[36] A. Sanderud and E. Sagstuen. EPR and ENDOR studies of single crystals of r-glycine X-ray irradiated at 295 K. *The Journal of Physical Chemistry B*, 102:9353–9361, 1998.

[37] S. Simon, M. Sodupe, and J. Bertrán. Isomerization versus fragmentation of glycine radical cation in gas phase. *The Journal of Physical Chemistry A*, 106:5697–5702, 2002.

[38] S. Kumar, A. K. Rai, V. Singh, and S. Rai. Vibrational spectrum of glycine molecule. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 61:2741–2746, 2005.

[39] J. Bertrán, L. Rodríguez-Santiago, and M. Sodupe. The different nature of bonding in Cu^{+}-glycine and Cu^{2+}-glycine. *The Journal of Physical Chemistry B*, 103:2310–2317, 1999.

[40] M. J. Locke, R. L. Hunter, and R. T. McIver. Experimental determination of the acidity and basicity of glycine in the gas phase. *Journal of the American Chemical Society*, 101:272–273, 1979.

[41] V. Barone, M. Biczysko, J. Bloino, and C. Puzzarini. Characterization of the elusive conformers of glycine from state-of-the-art structural, thermodynamic, and spectroscopic computations: Theory complements experiment. *Journal of Chemical Theory and Computation*, 9:1533–1547, 2013.

[42] M. L. Gordon, G. Cooper, C. Morin, T. Araki, C. C. Turci, K. Kaznatcheev, and A. P. Hitchcock. Inner-shell excitation spectroscopy of the peptide bond: Comparison of the c 1s, n 1 s, and o 1 s spectra of glycine, glycyl-glycine, and glycyl-glycyl-glycine. *The Journal of Physical Chemistry A*, 107:6144–6159, 2003.

[43] L. Rodríguez-Santiago, M. Sodupe, and J. Tortajada. Gas-phase reactivity of Ni+ with glycine. *The Journal of Physical Chemistry A*, 105:5340–5347, 2001.

[44] I. Corral, O. Mó, M. Yánez, J.-Y. Salpin, J. Tortajada, D. Moran, and L. Radom. An experimental and theoretical investigation of gas-phase reactions of Ca^{2+} with glycine. *Chemistry - A European Journal*, 12:6787–6796, 2006.

[45] S. Maclot, M. Capron, R. Maisonny, A. Ławicki, A. Méry, J. Rangama, J.-Y. Chesnel, S. Bari, R. Hoekstra, T. Schlathölter, B. Manil, L. Adoui, P. Rousseau, and B. A. Huber. Ion-induced fragmentation of amino acids: Effect of the environment. *A European Journal of Chemical Physics and Physical Chemistry*, 12:930–936, 2011.

[46] M. Szori, B. Jójárt, R. Izsák, K. Szori, I. G. Csizmadia, and B. Viskolcz. Chemical evolution of biomolecule building blocks. can thermodynamics explain the accumulation of glycine in the prebiotic ocean? *Physical Chemistry Chemical Physics*, 13:7449–7458, 2011.

[47] B. M. Messer, C. D. Cappa, J. D. Smith, K. R. Wilson, M. K. Gilles, R. C. Cohen, and R. J. Saykally. pH dependence of the electronic structure of glycine. *The Journal of Physical Chemistry B*, 109:5375–5382, 2005.

[48] P.-O. Sanskog, G. Nilsson, A. Lund, and T. Gilibro. For a pulse radiolysis and ESR spectroscopy study on ionized glycine in condensed phase. *The Journal of Physical Chemistry*, 85:2819, 1980.

[49] G. Depke, N. Heinrich, and H. Schwarz. On the gas phase chemistry of ionized glycine and its enol. A combined experimental and ab initio molecular orbital study. *International Journal of Mass Spectrometry and Ion Processes*, 62:99–117, 1984.

[50] C. V. Parast, K. K. Wong, S. A. Lewisch, J. W. Kozarich, J. Peisach, and R. S. Magliozzo. Hydrogen exchange of the glycyl radical of pyruvate formate-lyase is catalyzed by cysteine 419. *Biochemistry*, 34:2393–2399, 1995.

[51] P. Brunelle and A. Rauk. The radical model of Alzheimer's disease: Specific recognition of Gly29 and Gly33 by met35 in a β-sheet model of Abeta: An ONIOM study. *Journal of Alzheimer's Disease*, 4:283–289, 2002.

[52] V. Barone, M. Biczysko, and C. Puzzarini. Quantum chemistry meets spectroscopy for astrochemistry: Increasing complexity toward prebiotic molecules. *Accounts of Chemical Research*, 48:1413–1422, 2015.

[53] S. X. Tian. Conformation effects on the electronic structures of β-alanine. *The Journal of Physical Chemistry A*, 110:3961–3966, 2006.

[54] M. A. V. da Silva, M. d. D. M. C. Ribeiro da Silva, A. F. L. O. M. Santos, M. V. Roux, C. Foces-Foces, R. Notario, R. Guzmán-Mejía, and E. Juaristi. Experimental and computational thermochemical study of α-alanine (dl) and β-alanine. *The Journal of Physical Chemistry B*, 114:16471–16480, 2010.

[55] M. E. Sanz, A. Lesarri, M. I. Peña, V. Vaquero, V. Cortijo, J. C. López, and J. L. Alonso. The shape of β-alanine. *Journal of the American Chemical Society*, 128:3812–3817, 2006.

[56] L. Zhang, Y. Pan, H. Guo, T. Zhang, L. Sheng, F. Qi, P.-K. Lo, and K.-C. Lau. Conformation-specific pathways of β-alanine: a vacuum ultraviolet photoionization and theoretical study. *The Journal of Physical Chemistry A*, 113:5838–5845, 2009.

[57] O. I. Aruoma, M. J. Laughton, and B. Halliwell. Carnosine, homocarnosine and anserine: could they act as antioxidants in vivo? *The Biochemical journal*, 264:863–869, 1989.

[58] J. H. Kang, K. S. Kim, S. Y. Choi, H. Y. Kwon, M. H. Won, and T.-C. Kang. Protective effects of carnosine, homocarnosine and anserine against peroxyl radical-mediated Cu,Zn-superoxide dismutase modification. *Biochimica et Biophysica Acta (BBA)/General Subjects*, 1570:89–96, 2002.

[59] S. Y. Choi, H. Y. Kwon, O. B. Kwon, and J. H. Kang. Hydrogen peroxide-mediated Cu, Zn-superoxide dismutase fragmentation: protection by carnosine, homocarnosine and anserine. *Biochimica et Biophysica Acta (BBA)/General Subjects*, 1472:651–657, 1999.

[60] C. Renner, N. Zemitzsch, B. Fuchs, K. D. Geiger, M. Hermes, J. Hengstler, R. Gebhardt, J. Meixensberger, and F. Gaunitz. Carnosine retards tumor growth in vivo in an NIH3T3-HER2/neu mouse model. *Molecular Cancer*, 9:2, 2010.

[61] R. M. Hobson, B. Saunders, G. Ball, R. C. Harris, and C. Sale. Effects of β-alanine supplementation on exercise performance: a meta-analysis. *Amino Acids*, 43:25–37, 2012.

[62] G. G. Artioli, B. Gualano, A. Smith, J. Stout, and J. Lancha, Antonio Herbert. Role of β-alanine supplementation on muscle carnosine and exercise performance. *Medicine and Science in Sports and Exercise*, 42:1162–1173, 2010.

[63] A. S. Nutrition. β-alanine, fact sheets: Group b supplements, accessed: 04 of june, 2013. http://www.ausport.gov.au/ais/nutrition/supplements/group_b, 2009.

[64] E. Juaristi and V. Soloshonok, editors. *Enantioselective Synthesis of β-Amino Acids*. Wiley, 2005.

[65] F. S. Wu, T. T. Gibbs, and D. H. Farb. Dual activation of GABAA and glycine receptors by β-alanine: inverse modulation by progesterone and 5α-pregnan-3α-ol-20-one. *European Journal of Pharmacology*, 246:239–246, 1993.

[66] M. Sandberg and I. Jacobson. β-alanine, a possible neurotransmitter in the visual system? *Journal of Neurochemistry*, 37:1353–1356, 1981.

[67] V. Henkel, T. C. Baghai, D. Eser, P. Zill, R. Mergl, P. Zwanzger, C. Schüle, R. Bottlender, M. Jäger, R. Rupprecht, U. Hegerl, H.-J. Möller, and B. Bondy. The γ-amino butyric acid (GABA) receptor α-3 subunit gene polymorphism in unipolar depressive disorder: A genetic association study.

American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 126, 2004.

[68] B. Tighilet and M. Lacour. Gamma amino butyric acid (GABA) immunoreactivity in the vestibular nuclei of normal and unilateral vestibular neurectomized cats. *European Journal of Neuroscience*, 13, 2001.

[69] C. Cook. Monitoring on-line of extracellular γ-amino-4-butyric acid using microdialysis coupled to immunosensor analysis. *Journal of Neuroscience Methods*, 82, 1998.

[70] D. J. Garry, R. L. Sorenson, and H. D. Coulter. Ultrastructural localization of γ-amino butyric acid immunoreactivity in b cells of the rat pancreas. *Diabetologia*, 30, 1987.

[71] J. DIAMOND. Variation in the sensitivity to gamma-amino-butyric acid of different regions of the mauthner neurone. *Nature*, 199, 1963.

[72] A. Gil, S. Simon, L. Rodríguez-Santiago, J. Bertrán, and M. Sodupe. Influence of the Side Chain in the Structure and Fragmentation of Amino Acids Radical Cations. *Journal of Chemical Theory and Computation*, 3:2210–2220, 2007.

[73] A. Gil, S. Simon, M. Sodupe, and J. Bertrán. How the site of ionisation influences side-chain fragmentation in histidine radical cation. *Chemical Physics Letters*, 451:276–281, 2008.

[74] S. Naumov, I. Janovský, W. Knolle, and R. Mehnert. Radical cation, dimer radical cation and neutral radical of 2,3-dihydropyran - possible initiators of its polymerisation? *Macromolecular Chemistry and Physics*, 204:2099–2104, 2003.

[75] K. B. Nuzhdin, V. I. Feldman, and A. V. Kobzarenko. Delocalized methoxyacetone radical cation: structure and reactivity. *Mendeleev Communications*, 18:69–70, 2008.

[76] E. S. Shiryaeva, D. A. Tyurin, and V. I. Feldman. The structure and photochemical transformation of cyclopropylacetylene radical cation as revealed by matrix EPR and quantum chemical study. *Chemical Physics Letters*, 536:68–71, 2012.

[77] G. Marcotrigiano, L. Menabue, and G. C. Pellacani. Zinc(II), cadmium(II) and mercury(II) complexes with N-acetylglycine (aceturic acid) and their amine adducts. *Journal of Inorganic and Nuclear Chemistry*, 37, 1975.

[78] D. S. Martin, R. L. Stolfi, R. C. Sawyer, R. Nayak, S. Spiegelman, C. W. Young, and T. Woodcock. An overview of thymidine. *Cancer*, 45:1117–1128, 1980.

[79] K. V. Vidma, G. A. Bogdanchikov, A. V. Baklanov, D. A. Chestakov, and D. H. Parker. Experimental measurement of the van der waals binding energy of X-O2 clusters (X=Xe,CH3I,C3H6,C6H12). *The Journal of Chemical Physics*, 133, 2010.

[80] J. Camunas-Soler, A. Alemany, and F. Ritort. Experimental measurement of binding energy, selectivity, and allostery using fluctuation theorems. *Science*, 355:412–415, 2017.

[81] Y. Tsuchiya, T. Tamura, M. Fujii, and M. Ito. Keto-enol tautomer of uracil and thymine. *The Journal of Physical Chemistry*, 92:1760–1765, 1988.

[82] X. Hu, H. Li, W. Liang, and S. Han. Systematic study of the tautomerism of uracil induced by proton transfer. Exploration of water stabilization and mutagenicity. *The Journal of Physical Chemistry B*, 109:5935–5944, 2005.

[83] X. Hu, H. Li, L. Zhang, and S. Han. Tautomerism of uracil and 5-bromouracil in a microcosmic environment with water and metal ions. What roles do metal ions play? *The Journal of Physical Chemistry B*, 111:9347–9354, 2007.

[84] O. Brea, M. Yáñez, O. Mó, and A. M. Lamsabhi. On the stability of [(uracil)2-Cu]$^{2+}$ complexes in the gas phase. different pathways for the formation of [(uracil-h)(uracil)-Cu]+ monocations. *Organic & Biomolecular Chemistry*, 11:3862–3870, 2013.

[85] T. M. Miller, S. T. Arnold, A. A. Viggiano, and A. E. Stevens Miller. Acidity of a nucleotide base: Uracil. *The Journal of Physical Chemistry A*, 108:3439–3446, 2004.

[86] A. M. Lamsabhi, M. Alcamí, O. Mó, M. Yáñez, and J. Tortajada. Gas-phase deprotonation of uraci-Cu^{2+} and thiouracil-Cu^{2+} complexes. *The Journal of Physical Chemistry A*, 110:1943–1950, 2006.

[87] M. Tho Nguyen, A. K. Chandra, and Zeegers-Huyskens. Protonation and deprotonation energies of uracil. implications for the uracil-water complex. *Journal of the Chemical Society, Faraday Transactions*, 94:1277–1280, 1998.

[88] T. van Mourik, S. L. Price, and D. C. Clary. Ab initio calculations on uracil-water. *The Journal of Physical Chemistry A*, 103:1611–1618, 1999.

[89] E. S. Kryachko, M. T. Nguyen, and T. Zeegers-Huyskens. Theoretical study of uracil tautomers. 2. Interaction with water. *The Journal of Physical Chemistry A*, 105:1934–1943, 2001.

[90] J. Rejnek, M. Hanus, M. Kabelac, F. Ryjacek, and P. Hobza. Correlated ab initio study of nucleic acid bases and their tautomers in the gas phase, in a microhydrated environment and in aqueous solution. Part 4. Uracil and thymine. *Physical Chemistry Chemical Physics*, 7:2006–2017, 2005.

[91] S. Kim and H. F. Schaefer. Effects of microsolvation on the adenine-uracil base pair and its radical anion: Adenine-uracil mono- and dihydrates. *Journal of Physical Chemistry A*, 111, 2007.

[92] M. Capron, S. Díaz-Tendero, S. Maclot, A. Domaracka, E. Lattouf, A. Ławicki, R. Maisonny, J.-Y. Chesnel, A. Méry, J.-C. Poully, J. Rangama, L. Adoui, F. Martín, M. Alcamí, P. Rousseau, and B. A. Huber. A multicoincidence study of fragmentation dynamics in collision of γ-aminobutyric acid with low-energy ions. *Chemistry - A European Journal*, 18:9321–9332, 2012.

[93] D. Vernhet, J. Rozet, K. Wohrer, L. Adoui, C. Stéphan, A. Cassimi, and J. M. Ramillon. Excitation in swift heavy ion-atom collisions. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 107:71–78, 1996.

[94] B. J. Gaynor, R. G. Gilbert, and K. D. King. Intermolecular energy transfer in two-channel unimolecular reactions: the pyrolysis of 1-iodopropane. *Chemical Physics Letters*, 58:591–595, 1978.

[95] E. Guggenheim. Xlvi. on the determination of the velocity constant of a unimolecular reaction. *Philosophical Magazine Series 5*, 2:538–543, 1926.

[96] R. A. Jara-Toro, F. J. Hernández, R. A. Taccone, S. I. Lane, and G. A. Pino. Water catalysis of the reaction between methanol and OH at 294 K and the atmospheric implications. *Angewandte Chemie International Edition*, pages 2198–2202, 2017.

[97] S. Maclot, R. Delaunay, D. G. Piekarski, A. Domaracka, B. A. Huber, L. Adoui, F. Martín, M. Alcamí, L. Avaldi, P. Bolognesi, S. Díaz-Tendero, and P. Rousseau. Determination of energy-transfer distributions in ionizing ion-molecule collisions. *Physical Review Letters*, 117:073201, 2016.

[98] T. Schultz and M. Vrakking. *Attosecond and XUV Physics (Ultrafast Dynamics and Spectroscopy) // Multielectron High Harmonic Generation: Simple Man on a Complex Plane*, volume 10.1002/9783527677689, pages 201–256. Wiley-VCH Verlag GmbH and Co. KGaA, 2014.

[99] M. Born and J. R. Oppenheimer. On the quantum theory of molecules. *Annals of Physics*, 84:457, 1928.

[100] R. Gherib, L. Ye, I. G. Ryabinkin, and A. F. Izmaylov. On the inclusion of the diagonal Born-Oppenheimer correction in surface hopping methods. *The Journal of Chemical Physics*, 144:154103, 2016.

[101] M. Marinescu and A. Dalgarno. Long-range diagonal adiabatic corrections for the ground molecular state of alkali-metal dimers. *Physical Review A*, 57:1821–1826, 1998.

[102] V. Fock. jc slater. *Z. Physic*, 61:126, 1930.

[103] D. R. Hartree. The wave mechanics of an atom with a non-coulomb central field. Part I. Theory and methods. *Proceedings of the Cambridge Philosophical Society, Mathematical and physical sciences*, 24:89, 1928.

[104] C. Møller and P. M. S. Note on an approximation treatment for many-electron systems. *Physical Review*, 39:771, 1934.

[105] E. Condon. The theory of complex spectra. *Physical Review*, 36:1121, 1930.

[106] J. A. Pople, M. Head-Gordon, and K. Raghavachari. Quadratic configuration interaction. a general technique for determining electron correlation energies. *The Journal of chemical physics*, 87:5968–5975, 1987.

[107] J. Čížek. On the correlation problem in atomic and molecular systems. calculation of wavefunction components in ursell-type expansion using quantum-field theoretical methods. *The Journal of Chemical Physics*, 45:4256–4266, 1966.

[108] D. Bressanini, G. Morosi, and S. Tarasco. An investigation of nodal structures and the construction of trial wave functions. *The Journal of Chemical Physics*, 123:204109–204109–11, 2005.

[109] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Physical Review*, 136:B864–B871, 1964.

[110] R. G. Parr and Y. W. Density functional theory of atoms and molecules. *Oxford University Press*, 1989.

[111] T. Tsuneda. *Density functional theory in quantum chemistry*. Springer, 2014.

[112] E. Livshits and R. Baer. A well-tempered density functional theory of electrons in molecules. *Physical Chemistry Chemical Physics*, 9:2932–2941, 2007.

[113] T. Chachiyo. Communication: Simple and accurate uniform electron gas correlation energy for the full range of densities. *The Journal of Chemical Physics*, 145:021101, 2016.

[114] A. D. Becke. Density-functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics*, 98:5648–5652, 1993.

[115] Becke. Density functional exchange energy approximation with correct asymptotic behavior. *Physical Review A*, 38:3098–3100, 1988.

[116] B. Braïda, P. C. Hiberty, and A. Savin. A systematic failing of current density functionals: Overestimation of two-center three-electron bonding energies. *The Journal of Physical Chemistry A*, 102:7872–7877, 1998.

[117] Y. Zhao and D. G. Truhlar. The M06 suite of density functionals for main group thermochemistry,thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06 functionals and 12 other functionals. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)*, 119:525–525, 2008.

[118] D. G. Piekarski and S. Díaz-Tendero. Structure and stability of clusters of β-alanine in the gas phase: importance of the nature of intermolecular interactions. *Physical Chemistry Chemical Physics*, 19:5465–5476, 2017.

[119] P. Mori-Sánchez and A. J. Cohen. The derivative discontinuity of the exchange-correlation functional. *Physical Chemistry Chemical Physics*, 16:14378–14387, 2014.

[120] M. P. Andersson and P. Uvdal. New scale factors for harmonic vibrational frequencies using the b3lyp density functional method with the triple-χbasis set 6-311+g(d,p). *Journal of Physical Chemistry A*, 109:2937–2941, 2005.

[121] H. Jansen and P. Ros. Non-empirical molecular orbital calculations on the protonation of carbon monoxide. *Chemical Physics Letters*, 3, 1969.

[122] E. Miliordos and S. S. Xantheas. On the validity of the basis set superposition error and complete basis set limit extrapolations for the binding energy of the formic acid dimer. *The Journal of Chemical Physics*, 142, 2015.

[123] P. Rosmus. Ab initio methods in quantum chemistry, part I und part II. *Angewandte Chemie*, 99, 1987.

[124] P. Hobza and R. Zahradník. An essay on the theory and calculations of intermolecular interactions. *International Journal of Quantum Chemistry*, 42, 1992.

[125] D. B. Cook, J. A. Sordo, and T. L. Sordo. Some comments on the counterpoise correction for the basis set superposition error at the correlated level. *International Journal of Quantum Chemistry*, 48, 1993.

[126] S. Boys and F. Bernardi. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Molecular Physics*, 19:553–566, 1970.

[127] A. Johansson, P. Kollman, and S. Rothenberg. An application of the functional boys-bernardi counterpoise method to molecular potential surfaces. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)*, 29:167–172, 1973.

[128] J. P. Daudey, P. Claverie, and J. P. Malrieu. Perturbative ab initio calculations of intermolecular energies. i. method. *International Journal of Quantum Chemistry*, 8:1–15, 1974.

[129] D. J. Tannor. *Introduction to Quantum Mechanics: A Time-Dependent Perspective*. University Science Books, 2006.

[130] D. Marx and J. Hutter. *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*. Cambridge U. Press, 2009.

[131] C. E. Crespo-Hernández, B. Cohen, P. M. Hare, and B. Kohler. Ultrafast excited-state dynamics in nucleic acids. *Chemical reviews*, 104:1977–2020, 2004.

[132] E. Tapavicza, I. Tavernelli, and U. Rothlisberger. Trajectory surface hopping within linear response time-dependent density-functional theory. *Physical Review Letters*, 98:023001, 2007.

[133] M. Richter, P. Marquetand, J. González-Vázquez, I. Sola, and L. González. Sharc: ab initio molecular dynamics with surface hopping in the adiabatic representation including arbitrary couplings. *Journal of Chemical Theory and Computation*, 7:1253–1258, 2011.

[134] A. Dreuw and M. Head-Gordon. Single-reference ab initio methods for the calculation of excited states of large molecules. *Chemical Reviews*, 105:4009–4037, 2005.

[135] A. Muzas, F. Martín, and C. Díaz. Scattering of H (D) from LiF (100) under fast grazing incidence conditions: to what extent is classical dynamics a useful tool? *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 354:9–15, 2015.

[136] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. J. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. A second generation force field for the simulation of proteins, nucleic acids and organic molecules. *Journal of the American Chemical Society*, 117:5179–5197, 1995.

[137] P. K. Weiner and P. A. Kollman. Amber: Assisted model building with energy refinement. a general program for modeling molecules and their interactions. *Journal of Computational Chemistry*, 2:287–303, 1981.

[138] S. Weiner, P. Kollman, D. Case, U. Singh, C. Ghio, G. Alagona, S. J. Profeta, and P. Weiner. A new force field for molecular mechanical simulation of nucleic acids and proteins. *Journal of the American Chemical Society*, 106:765–784, 1984.

[139] S. J. Weiner, P. A. Kollman, D. T. Nguyen, and D. A. Case. An all atom force field for simulations of proteins and nucleic acids. *Journal of Computational Chemistry*, 7:230–252, 1986.

[140] TURBOMOLE V7.0 2015, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from
`http://www.turbomole.com`.

[141] H. B. Schlegel, J. M. Millam, S. S. Iyengar, G. A. Voth, A. D. Daniels, G. E. Scuseria, and M. J. Frisch. Ab initio molecular dynamics: Propagating the density matrix with Gaussian orbitals. *The Journal of Chemical Physics*, 114:9758–9763, 2001.

[142] H. B. Schlegel, S. S. Iyengar, X. Li, J. M. Millam, G. A. Voth, G. E. Scuseria, and M. J. Frisch. Ab initio molecular dynamics: Propagating the density matrix with Gaussian orbitals. III. Comparison with Born–Oppenheimer dynamics. *The Journal of Chemical Physics*, 117:8694–8704, 2002.

[143] R. Car and M. Parrinello. Unified approach for molecular dynamics and density-functional theory. *Physical Review Letters*, 55:2471–2474, 1985.

[144] L. Verlet. Computer "experiments" on classical fluids. i. thermodynamical properties of lennard-jones molecules. *Physical Review (Series I)*, 159:98–103, 1967.

[145] L. Verlet. Computer "experiments" on classical fluids. II. Equilibrium correlation functions. *Physical Review (Series I)*, 165:201–214, 1968.

[146] W. F. V. Gunsteren and H. J. C. Berendsen. A leap-frog algorithm for stochastic dynamics. *Molecular Simulation*, 1:173–185, 1988.

[147] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case. Development and testing of a general amber force field. *Journal of Computational Chemistry*, 25:1157–1174, 2004.

[148] W. D. Cornell, P. Cieplak, C. I. Bayly, and P. A. Kollmann. Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation. *Journal of the American Chemical Society*, 115:9620–9631, 1993.

[149] S. S. Iyengar, H. B. Schlegel, J. M. Millam, G. A. Voth, G. E. Scuseria, and M. J. Frisch. Ab initio molecular dynamics: Propagating the density matrix with Gaussian orbitals. II. Generalizations based on mass-weighting, idempotency, energy conservation and choice of initial conditions. *The Journal of Chemical Physics*, 115:10291–10302, 2001.

[150] S. S. Iyengar, H. B. Schlegel, and G. A. Voth. Atom-centered density matrix propagation (ADMP): Generalizations using bohmian mechanics. *The Journal of Physical Chemistry A*, 107:7269–7277, 2003.

[151] S. Goedecker. Linear scaling electronic structure methods. *Reviews of Modern Physics*, 71:1085–1123, 1999.

[152] S. G. E. Linear scaling density functional calculations with Gaussian orbitals. *The Journal of Physical Chemistry A*, 104:4782–4790, 1999.

[153] M. K. Cyrański, T. M. Krygowski, M. Wisiorowski, N. J. R. van Eikema Hommes, and P. v. R. Schleyer. Global and local aromaticity in porphyrins: An analysis based on molecular geometries and nucleus-independent chemical shifts. *Angewandte Chemie International Edition*, 37:177–80, 1998.

[154] F. Weinhold and C. R. Landis. *Discovering chemistry with natural bond orbitals*. Wiley, 2012.

[155] J. P. Foster and F. Weinhold. Natural hybrid orbitals. *Journal of the American Chemical Society*, 102:7211–7218, 1980.

[156] R. F. W. Bader. *Atoms in Molecules: A Quantum Theory*, volume l. Clarendon Press: Oxford, England, 1990.

[157] R. F. W. Bader and P. M. Beddall. Virial field relationship for molecular charge distributions and the spatial partitioning of molecular properties. *The Journal of Chemical Physics*, 56:3320–3329, 1972.

[158] T. Koopmans. Über die zuordnung von wellenfunktionen und eigenwerten zu den einzelnen elektronen eines atoms. *Physica*, 1:104 – 113, 1934.

[159] L. S. Cederbaum. Direct calculation of ionization potentials of closed-shell atoms and molecules. *Theoretica Chimica Acta*, 31:239–260, 1973.

[160] L. S. Cederbaum. One-body green's function for atoms and molecules: theory and application. *Journal of Physics B Atomic and Molecular Physics*, 8:290–303, 1975.

[161] I. Prigogine and S. A. Rice. *Advances in Chemical Physics Advances in Chemical Physics*, volume 36, chapter Theoretical Aspects of Ionization Potentials and Photoelectron Spectroscopy: A Green's Function Approach, pages 205–344. John Wiley and Sons, Inc., 1977.

[162] G. A. Heal, P. D. Walker, P. G. Mezey, and M. Ramek. Shape-similarity analysis of 20 stable conformations of neutral β-alanine. *Canadian Journal of Chemistry*, 74:1660–1670, 1996.

[163] M. Ramek. Ab initio SCF investigation of β-alanine. *Journal of Molecular Structure: THEOCHEM*, 208:301, 1990.

[164] M. Ramek and V. K. W. Cheng. On the role of polarization functions in scf calculations of glycine and related systems with intramolecular hydrogen bonding. *International Journal of Quantum Chemistry*, 44:15–26, 1992.

[165] K. Fukui. The path of chemical-reactions - the IRC approach. *Accounts of Chemical Research*, 14:363–368, 1981.

[166] C. Peng and H. B. Schlegel. Combining synchronous transit and quasi-newton methods for finding transition states. *Israel Journal of Chemistry*, 33:449–54, 1994.

[167] C. Peng, P. Y. Ayala, H. B. Schlegel, and M. J. Frisch. Using redundant internal coordinates to optimize equilibrium geometries and transition states. *Journal of Computational Chemistry*, 17:49–56, 1996.

[168] D. A. Gibson, I. V. Ionova, and E. A. Carter. A comparison of Car-Parrinello and Born-Oppenheimer generalized valence bond molecular dynamics. *Chemical Physics Letters*, 240:261–267, 1995.

[169] S. Simon, A. Gil, M. Sodupe, and J. Bertrán. Structure and fragmentation of glycine, alanine, serine and cysteine radical cations. A theoretical study. *Journal of Molecular Structure: THEOCHEM*, 727:191–197, 2005.

[170] R. de Nalda and L. Bañares, editors. *Ultrafast Phenomena in Molecular Sciences*, volume 107 of *Springer Series in Chemical Physics*. Springer International Publishing, Heidelberg New York Dordrecht London, 2014.

[171] F. Calegari, D. Ayuso, A. Trabattoni, L. Belshaw, S. De Camillis, S. Anumula, F. Frassetto, L. Poletto, A. Palacios, P. Decleva, J. B. Greenwood, F. Martín, and M. Nisoli. Ultrafast electron dynamics in phenylalanine initiated by attosecond pulses. *Science*, 346:336–339, 2014.

[172] B. Erk, D. Rolles, L. Foucar, B. Rudek, S. W. Epp, M. Cryle, C. Bostedt, S. Schorb, J. Bozek, A. Rouzee, A. Hundertmark, T. Marchenko, M. Simon, F. Filsinger, L. Christensen, S. De, S. Trippel, J. Küpper, H. Stapelfeldt, S. Wada, K. Ueda, M. Swiggers, M. Messerschmidt, C. D. Schröter, R. Moshammer, I. Schlichting, J. Ullrich, and A. Rudenko. Ultrafast charge rearrangement and nuclear dynamics upon inner-shell multiple ionization of small polyatomic molecules. *Physical Review Letters*, 110:053003, 2013.

[173] A. N. Markevitch, D. A. Romanov, S. M. Smith, and R. J. Levis. Coulomb explosion of large polyatomic molecules assisted by nonadiabatic charge localization. *Physical Review Letters*, 92:063001, 2004.

[174] Y. H. Jiang, A. Rudenko, O. Herrwerth, L. Foucar, M. Kurka, K. U. Kühnel, M. Lezius, M. F. Kling, J. van Tilborg, A. Belkacem, K. Ueda, S. Düsterer, R. Treusch, C. D. Schröter, R. Moshammer, and J. Ullrich. Ultrafast extreme ultraviolet induced isomerization of acetylene cations. *Physical Review Letters*, 105:263002, 2010.

[175] S. De, J. Rajput, A. Roy, P. N. Ghosh, and C. P. Safvan. Formation of H_3^+ due to intramolecular bond rearrangement in doubly charged methanol. *Physical Review Letters*, 97:213201, 2006.

[176] S. Maclot, D. G. Piekarski, A. Domaracka, A. Méry, V. Vizcaino, L. Adoui, F. Martín, M. Alcamí, B. A. Huber, P. Rousseau, and S. Díaz-Tendero. Dynamics of glycine dications in the gas phase:ultrafast intramolecular hydrogen migration versus coulomb repulsion. *The Journal of Physical Chemistry Letters*, 4:3903–3909, 2013.

[177] D. G. Piekarski, R. Delaunay, S. Maclot, L. Adoui, F. Martín, M. Alcamí, B. A. Huber, P. Rousseau, A. Domaracka, and S. Díaz-Tendero. Unusual hydroxyl migration in the fragmentation of β-alanine dication in the gas phase. *Physical Chemistry Chemical Physics*, 17:16767–16778, 2015.

[178] L. Levin, W. Skomorowski, L. Rybak, R. Kosloff, C. P. Koch, and Z. Amitay. Coherent control of bond making. *Physical Review Letters*, 114:233003, 2015.

[179] Z. Li, O. Vendrell, and R. Santra. Ultrafast charge transfer of a valence double hole in glycine driven exclusively by nuclear motion. *Physical Review Letters*, 115:143002, 2015.

[180] L. Chen, S. Martin, J. Bernard, and R. Brédy. Direct measurement of internal energy of fragmented C60. *Physical Review Letters*, 98:193401, 2007.

[181] R. Brédy, J. Bernard, L. Chen, G. Montagne, B. Li, and S. Martin. Fragmentation of adenine under energy control. *The Journal of Chemical Physics*, 130:114305, 2009.

[182] M. Durante and J. S. Loeffler. Charged particles in radiation oncology. *Nature Reviews Clinical Oncology*, 7:37–43, 2010.

[183] A. V. Solov'yov, E. Surdutovich, E. Scifoni, I. Mishustin, and W. Greiner. Physics of ion beam cancer therapy: A multiscale approach. *Physical Review E*, 79:011909, 2009.

[184] B. Liu, S. Nielsen, P. Hvelplund, H. Zettergren, H. Cederquist, B. Manil, and B. A. Huber. Collision-Induced Dissociation of Hydrated Adenosine Monophosphate Nucleotide Ions: Protection of the Ion in Water Nanoclusters. *Physical Review Letters*, 97:133401, 2006.

[185] S. Ptasińska, S. Denifl, V. Grill, T. Märk, E. Illenberger, and P. Scheier. Bond- and Site-Selective Loss of H- from Pyrimidine Bases. *Physical Review Letters*, 95:093201, 2005.

[186] E. C. Montenegro, M. B. Shah, H. Luna, S. W. J. Scully, A. L. F. de Barros, J. A. Wyer, and J. Lecointre. Water fragmentation and energy loss by carbon ions at the distal region of the bragg peak. *Physical Review Letters*, 99:213201, 2007.

[187] H. Levola, K. Kooser, E. Rachlew, E. Nõmmiste, and E. Kukk. Fragmentation of thymidine induced by ultraviolet photoionization and thermal degradation. *Int. Journal of Mass Spectrometry*, 353:7–11, 2013.

[188] T. Bergen, X. Biquard, A. Brenac, F. Chandezon, B. A. Huber, D. Jalabert, H. Lebius, M. Maurel, E. Monnand, J. Opitz, A. Pesnelle, B. Pras, C. Ristori, and J. C. Rocco. Multiply charged cluster ion crossed-beam apparatus: Multi-ionization of clusters by ion impact. *Review of Scientific Instruments*, 70:3244–3253, 1999.

[189] O. Plekan, M. Coreno, V. Feyer, A. Moise, R. Richter, M. de Simone, R. Sankari, and K. C. Prince. Electronic state resolved pepico spectroscopy of pyrimidine. *Physica Scripta*, 78:058105, 2008.

[190] E. Itälä, M. A. Huels, E. Rachlew, K. Kooser, T. Hägerth, and E. Kukk. A comparative study of dissociation of thymidine molecules following valence or core photoionization. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 46:215102, 2013.

[191] Z. Deng, I. Bald, E. Illenberger, and M. Huels. Beyond the bragg peak: Hyperthermal heavy ion damage to DNA components. *Physical Review Letters*, 95:153201, 2005.

[192] S. Ptasińska, P. Candori, S. Denifl, S. Yoon, V. Grill, P. Scheier, and T. Märk. Dissociative ionization of the nucleosides thymidine and uridine by electron impact. *Chemical Physics Letters*, 409:270–276, 2005.

[193] G. Martinet, S. Díaz-Tendero, M. Chabot, K. Wohrer, S. D. Negra, F. Mezdari, H. Hamrita, P. Désesquelles, A. L. Padellec, D. Gardés, L. Lavergne, G. Lalu, X. Grave, J. F. Clavelin, P.-A. Hervieux, M. Alcamí, and F. Martín. Fragmentation of highly excited small neutral carbon clusters. *Physical Review Letters*, 93:063401, 2004.

[194] P. Bolognesi, J. A. Kettunen, A. Cartoni, R. Richter, S. Tosic, S. Maclot, P. Rousseau, R. Delaunay, and L. Avaldi. Site- and state-selected photofragmentation of 2br-pyrimidine. *Physical Chemistry Chemical Physics*, 17:24063–24069, 2015.

[195] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken,

C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox. Gaussian 09 Revision B.01, 2010. Gaussian Inc. Wallingford CT 2009.

[196] D. L. Yeager. *Multiconfigurational Green's function (propagator) techniques for excitation energies, ionization potentials and electron affinities: an overview*. Springer US, New York, d. mukherjee edition, 1992.

[197] M. A. Hervé Du Penhoat, P. López-Tarifa, K. K. Ghose, Y. Jeanvoine, M. P. Gaigeot, R. Vuilleumier, M. F. Politis, and M. C. Bacchus-Montabonel. Modeling proton-induced damage on 2-deoxy-d-ribose. conformational analysis. *Journal of Molecular Modeling*, 20:2221, 2014.

[198] A. Rentenier, L. F. Ruiz, S. Díaz-Tendero, B. Zarour, P. Moretto-Capelle, D. Bordenave-Montesquieu, A. Bordenave-Montesquieu, P.-A. Hervieux, M. Alcamí, M. F. Politis, J. Hanssen, and F. Martín. Absolute charge transfer and fragmentation cross sections in He^{2+}-C_{60} collisions. *Physical Review Letters*, 100:183401, 2008.

[199] A. Bárány, G. Astner, H. Cederquist, H. Danared, S. Huldt, P. Hvelplund, A. Johnson, H. Knudsen, L. Liljeby, and K.-G. Rensfelt. Absolute cross sections for multi-electron processes in low energy Ar^{q+}-Ar collisions: Comparison with theory. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 9:397 – 399, 1985.

[200] S. Díaz-Tendero, L. F. Ruiz, B. Zarour, J. Hanssen, M. Alcamí, M. F. Politis, P.-A. Hervieux, and F. Martín. Absolute charge transfer and fragmentation cross sections in He^{2+}+C_{60} collisions. *Journal of Physics: Conference Series*, 194:012047, 2009.

[201] S. Bari, R. Hoekstra, and T. Schlathölter. Peptide fragmentation by keV ion-induced dissociation. *Physical Chemistry Chemical Physics*, 12:3376–3383, 2010.

[202] M.-C. Bacchus-Montabonel. Theoretical study of charge transfer dynamics in collisions of C^{6+} carbon ions with pyrimidine nucleobases. *The European Physical Journal D*, 66:1–7, 2012.

[203] J.-P. Champeaux, P. Çarçabal, J. Rabier, P. Cafarelli, M. Sence, and P. Moretto-Capelle. Dehalogenation of 5-halo-uracil molecules induced by 100 keV proton collisions. *Physical Chemistry Chemical Physics*, 12:5454–5461, 2010.

[204] D. Almeida, D. Kinzel, F. Ferreira da Silva, B. Puschnigg, D. Gschliesser, P. Scheier, S. Denifl, G. Garcia, L. Gonzalez, and P. Limao-Vieira. N-site de-methylation in pyrimidine bases as studied by low energy electrons and ab initio calculations. *Physical Chemistry Chemical Physics*, 15:11431–11440, 2013.

[205] F. Alvarado, J. Bernard, B. Li, R. Brédy, L. Chen, R. Hoekstra, S. Martin, and T. Schlathölter. Precise determination of 2-deoxy-d-ribose internal energies after keV proton collisions. *A European Journal of Chemical Physics and Physical Chemistry*, 9:1254–1258, 2008.

[206] M. H. Holzscheiter, N. Bassler, M. Dosanjh, B. S. Sørensen, and J. Overgaard. A community call for a dedicated radiobiological research facility to support particle beam cancer therapy. *Radiotherapy and Oncology*, 105:1–3, 2012.

[207] P. Swiderek. Fundamental processes in radiation damage of DNA. *Angewandte Chemie International Edition*, 45:4056–4059, 2006.

[208] V. Eulenburg, W. Armsen, H. Betz, and J. Gomeza. Glycine transporters: essential regulators of neurotransmission. *Trends in Biochemical Sciences*, 30:325–333, 2005.

[209] G. Junk and H. Svec. The mass spectra of the α-amino acids. *Journal of the American Chemical Society*, 85:839–845, 1963.

[210] H.-W. Jochims, M. Schwell, J.-L. Chotin, M. Clemino, F. Dulieu, H. Baumgärtel, and S. Leach. Photoion mass spectrometry of five amino acids in the 6-22 eV photon energy range. *Chemical Physics*, 298:279–297, 2004.

[211] S. Bari, P. Sobociński, J. Postma, F. Alvarado, R. Hoekstra, V. Bernigaud, B. Manil, J. Rangama, B. Huber, and T. Schlathölter. Fragmentation of α- and β-alanine molecules by ions at Bragg-peak energies. *The Journal of Chemical Physics*, 128:074306, 2008.

[212] H. Zettergren, P. Rousseau, Y. Wang, F. Seitz, T. Chen, M. Gatchell, J. D. Alexander, M. H. Stockett, J. Rangama, J. Y. Chesnel, M. Capron, J. C. Poully, A. Domaracka, A. Méry, S. Maclot, H. T. Schmidt, L. Adoui, M. Alcamí, A. G. G. M. Tielens, F. Martín, B. A. Huber, and H. Cederquist. Formations of dumbbell C_{118} and C_{119} inside clusters of C_{60} molecules by collision with α particles. *Physical Review Letters*, 110:185501, 2013.

[213] F. Chandezon, B. Huber, and C. Ristori. A new-regime Wiley-McLaren time-of-flight mass spectrometer. *Review of Scientific Instruments*, 65:3344–3353, 1994.

[214] V. Bernigaud, O. Kamalou, A. Lawicki, M. Capron, R. Maisonny, B. Manil, L. Maunoury, J. Rangama, P. Rousseau, J.-Y. Chesnel, L. Adoui, and B. A. Huber. ARIBE: a low energy ion beam facility in Caen. *Publications de l'Observatoire Astronomique de Beograd*, 84:83–86, 2008.

[215] C. Lee, W. Yang, and R. G. Parr. Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Physical Review B*, 37:785–789, 1988.

[216] A. G. C. Conformers of gaseous glycine. *Journal of the American Chemical Society*, 114:9568–9575, 1992.

[217] J. Daligault, F. Chandezon, C. Guet, B. A. Huber, and S. Tomita. Energy transfer in collisions of metal clusters with multiply charged ions. *Physical Review A*, 66:033205, 2002.

[218] P. Rousseau, A. Ławicki, A. I. S. Holm, M. Capron, R. Maisonny, S. Maclot, E. Lattouf, H. A. B. Johansson, F. Seitz, A. Méry, J. Rangama, H. Zettergren, S. Rosén, H. T. Schmidt, J.-Y. Chesnel, A. Domaracka, B. Manil, L. Adoui, H. Cederquist, and B. A. Huber. Low-energy ions interacting with anthracene molecules and clusters. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 279:140–143, 2012.

[219] A. Niehaus. A classical model for multiple-electron capture in slow collisions of highly charged ions with atoms. *Journal of Physics B: Atomic and Molecular Physics*, 19:2925, 1986.

[220] COST Action MP1002: Nano-IBCT: Nano-scale insight into ion beam cancer therapy, 2010.

[221] G. A. Olah, A. Germain, H. C. Lin, and D. A. Forsyth. Electrophilic reactions at single bonds. XVIII. Indication of protosolvated de facto substituting agents in the reactions of alkanes with acetylium and nitronium ions in superacidic media. *Journal of the American Chemical Society*, 97:2928–2929, 1975.

[222] G. A. Olah. Superelectrophiles. *Angewandte Chemie International Edition*, 32:767–788, 1993.

[223] L. Belshaw, F. Calegari, M. J. Duffy, A. Trabattoni, L. Poletto, M. Nisoli, and J. B. Greenwood. Observation of ultrafast charge migration in an amino acid. *The Journal of Physical Chemistry Letters*, 3:3751–3754, 2012.

[224] M. E.-A. Madjet, O. Vendrell, and R. Santra. Ultrafast dynamics of photoionized acetylene. *Physical Review Letters*, 107:263002, 2011.

[225] R. Itakura, P. Liu, Y. Furukawa, T. Okino, K. Yamanouchi, and H. Nakano. Two-body coulomb explosion and hydrogen migration in methanol induced by intense 7 and 21 fs laser pulses. *The Journal of Chemical Physics*, 127:104306, 2007.

[226] O. González-Magaña, M. Tiemens, G. Reitsma, L. Boschman, M. Door, S. Bari, P. O. Lahaie, J. R. Wagner, M. A. Huels, R. Hoekstra, and T. Schlathölter. Fragmentation of protonated oligonucleotides by energetic photons and C^{q+} ions. *Physical Review A*, 87:032702, 2013.

[227] B. Jochim, A. Lueking, L. Doshier, S. Carey, E. Wells, E. Parke, M. Leonard, K. D. Carnes, and I. Ben-Itzhak. Rapid formation of H^{3+} from ammonia and methane following 4 MeV proton impact. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 42:091002, 2009.

[228] S. Bari, F. Alvarado, J. Postma, P. Sobociński, R. Hoekstra, and T. Schlathölter. Kinetic energy releases of small amino acids upon interaction with keV ions. *The European Physical Journal D*, 51:81–87, 2008.

[229] J. Tamuliene, L. Romanova, V. Vukstich, and A. Snegursky. Mechanisms of the electron-impact-induced glycine molecule fragmentation. *Chemical Physics*, 404:36 – 41, 2012.

[230] M. J. Polce and C. Wesdemiotis. First generation and characterization of the enol of glycine, H_2N-CH-$C(OH)_2$, in the gas phase. *Journal of Mass Spectrometry*, 35:251–257, 2000.

[231] T. Marino and N. Russo. On the fragmentation pathway of the ionized enol of glycine in the gas phase. *Rap. Comm. Mass Spectrom.*, 15:541–545, 2001.

[232] J. E. Dickens, W. M. Irvine, C. H. DeVries, and M. Ohishi. Hydrogenation of interstellar molecules: A survey for methylenimine (CH_2NH). *The Astrophysical Journal*, 479:307, 1997.

[233] A. J. Remijan, J. M. Hollis, F. J. Lovas, W. D. Stork, P. R. Jewell, and D. S. Meier. Detection of interstellar cyanoformaldehyde (CNCHO). *The Astrophysical Journal Letters*, 675:85, 2008.

[234] R. N. Loeppky and C. J. Michejda. Nitrosamines and related N-nitroso compounds. *ACS Symposium Series*, 553:1–18, 1994.

[235] J. Casado, A. Castro, J. R. Leis, M. Mosquera, and M. E. Pena. The mechanism of the nitrosation of α-amino acids: evidence for an intramolecular pathway. *Journal of the Chemical Society, Perkin Transactions 2*, 12:1859–1864, 1985.

[236] R. Gil, J. Casado, and C. Izquierdo. The nitrosation of amino acids. *International Journal of Chemical Kinetics*, 26:1167–1178, 1994.

[237] G. A. Olah, J. Shih, and G. K. S. Prakash. Preparation of α-bromo- and α-chlorocarboxylic acids from α-amino acids. *Helvetica Chimica Acta*, 66:1028–1030, 1983.

[238] S. Bari, F. Alvarado, J. Postma, P. Sobociński, R. Hoekstra, and T. Schlathölter. Kinetic energy releases of small amino acids upon interaction with keV ions. *The European Physical Journal D*, 51:81–87, 2008.

[239] P. Sobociński, S. Bari, J. Postma, F. Alvarado, R. Hoekstra, B. Manil, J. Rangama, V. Bernigaud, B. Huber, and T. Schlathölter. Isomeric effects in ion-induced fragmentation of α- and β-alanine. *Journal of Physics: Conference Series*, 101:012006, 2008.

[240] M. Dizdaroglu. Chemical determination of free radical-induced damage to DNA. *Free Radical Biology and Medicine*, 10:225–242, 1991.

[241] P. Riley. Free radicals in biology: Oxidative stress and the effects of ionizing radiation. *International Journal of Radiation Biology*, 65:27–33, 1994.

[242] J. Cadet, T. Douki, and J.-L. Ravanat. Oxidatively generated base damage to cellular DNA. *Free Radical Biology and Medicine*, 49:9–21, 2010.

[243] S. S. Wallace. Supplement: Madame Curie's discovery of Radium (1898): A commemoration by women in radiation sciences || Enzymatic processing of radiation-induced free radical damage in DNA. *Radiation Research*, 150:S60–S79, 1998.

[244] B. Boudaïffa, P. Cloutier, D. Hunting, M. Huels, and S. L. Resonant formation of DNA strand breaks by low-energy (3 to 20 eV) electrons. *Science*, 287:1658–1660, 2000.

[245] P. Bolognesi, P. O'Keeffe, and L. Avaldi. Soft x-ray interaction with organic molecules of biological interest. In G. García Gómez-Tejedor and M. C. Fuss, editors, *Radiation Damage in Biomolecular Systems*, Biological and Medical Physics, Biomedical Engineering, pages 165–176. Springer Netherlands, 2012.

[246] T. Toraya. Radical catalysis in coenzyme B12-dependent isomerization (eliminating) reactions. *Chemical Reviews*, 34:2095–2127, 2003.

[247] K. L. Brown. Chemistry and enzymology of vitamin B12. *Chemical Reviews*, 36:2075–2149, 2005.

[248] D. M. Smith, B. T. Golding, and L. Radom. Understanding the mechanism of B12-dependent diol dehydratase: A synergistic retro-push-pull proposal. *Journal of the American Chemical Society*, 123:1664–1675, 2001.

[249] T. Toraya, K. Yoshizawa, and T. Kamachi. Catalytic roles of active-site amino acid residues of coenzyme B12-dependent diol dehydratase: Protonation state of histidine and pull effect of glutamate. *Journal of the American Chemical Society*, 126:16207–16216, 2004.

[250] T. Kamachi, T. Toraya, and K. Yoshizawa. Computational mutation analysis of hydrogen abstraction and radical rearrangement steps in the catalysis of coenzyme B12-dependent diol dehydratase. *Chemistry - A European Journal*, 13:7864–7873, 2007.

[251] S. Maclot, D. G. Piekarski, R. Delaunay, A. Domaracka, A. Méry, V. Vizcaino, J.-Y. Chesnel, F. Martín, M. Alcamí, B. A. Huber, L. Adoui, P. Rousseau, and S. Díaz-Tendero. Stability of the glycine cation in the gas phase after interaction with multiply charged ions. *The European Physical Journal D - Atomic, Molecular, Optical and Plasma Physics*, 68:149, 2014.

[252] J. L. Holmes and T. S. Jean. The mass spectra of carboxylic acids-II: Fragmentation mechanisms in the homologous series $HOOC(CH_2)COOH$. *Journal of Mass Spectrometry*, 3:1505–1518, 1970.

[253] A. G. Harrison. To b or not to b: the ongoing saga of peptide b ions. *Mass Spectrometry Reviews*, 28:640–654, 2009.

[254] G. A. Olah, G. K. S. Prakash, and Y. L. Chao. Synthetic methods and reactions. Part 106. Suppression of anchimerically assisted rearrangement products in the synthesis of α-fluorocarboxylic acids from α-amino acids with 48:52 (w/w) hydrogen fluoride/pyridine [1]. *Helvetica Chimica Acta*, 64:2528–2530, 1981.

[255] A. I. Kuleff and L. S. Cederbaum. Charge migration in different conformers of glycine: The role of nuclear geometry. *Chemical Physics*, 338:320–328, 2007.

[256] J. Andrés, L. R. Domingo, M. T. Picher, and V. S. Safont. Comparative theoretical study of transition structures, barrier heights, and reaction energies for the intramolecular tautomerization in acetaldehyde/vinyl alcohol and acetaldimine/vinylamine systems. *International Journal of Quantum Chemistry*, 66:9–24, 1998.

[257] J. Eland, F. Wort, and R. Royds. A photoelectron-ion-ion triple coincidence technique for the study of double photoionization and its consequences. *Journal of Electron Spectroscopy and Related Phenomena*, 41:297–309, 1986.

[258] V. E. Selezneva, A. I. Chmarin, T. L. Golitsyna, and D. S. Zhuk. Polymerization characteristics of amino derivatives of ethylenimine. *Russian Chemical Bulletin*, 24:189–189, 1975.

[259] K. E. Kruglyakova, D. L. Zybina, B.Ulanov, and N. M. Emanuel. The kinetic characteristics of the action of chemical mutagens (ethylenimine derivatives) on DNA. *Doklady Akademii Nauk SSSR*, 718:161, 1965.

[260] S. Naumov, I. Janovský, W. Knolle, R. Mehnert, and D. A. Turin. Low-temperature EPR and quantum chemical study of lactone radical cations and their transformations. *Radiation Physics and Chemistry*, 73:206–212, 2005.

[261] M. A. Freitas, R. A. J. O'Hair, J. A. R. Schmidta, S. E. Tichy, B. E. Plashko, and T. D. Williams. Gas-phase reactions of glycine, alanine, valine and their N-methyl derivatives with the nitrosonium ion, NO^+. *Journal of Mass Spectrometry*, 31:1086–1092, 1996.

[262] M. C. Blanchette, J. L. Holmes, C. E. C. A. Hop, F. P. Lossing, R. Postma, P. J. A. Ruttink, and J. K. Terlouw. Theory and experiment in concert: the $[MeOC{:}O]^+$ ion and its isomers. *Journal of the American Chemical Society*, 108:7589–7594, 1986.

[263] M. H. Lien and A. C. Hopkinson. A theoretical study of α-acylmethyl, oxiranyl, and acetyl carbocations. *Journal of the American Chemical Society*, 110:3788–3792, 1988.

[264] D. Schröder, N. Goldberg, W. Zummack, H. Schwarz, J. C. Poutsma, and R. R. Squires. Generation of α-acetolactone and the acetoxyl diradical CH_2COO in the gas phase. *International Journal of Mass Spectrometry and Ion Processes*, 165-166:71–82, 1997.

[265] A. Cruz-Torres, A. Galano, and J. R. Alvarez-Idaboy. Kinetics and mechanism of the β-alanine + OH gas phase reaction: a quantum mechanical approach. *Physical Chemistry Chemical Physics*, 8:285–292, 2006.

[266] S. Kwok. Organic matter in space: from star dust to the solar system. *Astrophysics and Space Science*, 319:5–21, 2009.

[267] T. Wang and J. H. Bowie. Radical routes to interstellar glycolaldehyde. The possibility of stereoselectivity in gas-phase polymerization reactions involving CH_2O and CH_2OH. *Organic and Biomolecular Chemistry*, 8:4757–0, 2010.

[268] J. E. Bartmess. Gas-phase equilibrium affinity scales and chemical ionization mass spectrometry. *Mass Spectrometry Reviews*, 8:297–343, 1989.

[269] B. Halliwell and J. M. C. Gutteridge. *Free Radicals in Biology and Medicine*. Oxford University Press, Oxford, 5th edition, 2015.

[270] J. Cadet and J. R. Wagner. DNA base damage by reactive oxygen species, oxidizing agents, and UV radiation. *Cold Spring Harbor Perspectives in Biology*, 5:a012559, 2013.

[271] R. A. Mitteer, Y. Wang, J. Shah, S. Gordon, M. Fager, P.-P. Butter, H. J. Kim, C. Guardiola-Salmeron, A. Carabe-Fernandez, and Y. Fan. Proton beam radiation induces DNA damage and cell apoptosis in glioma stem cells through reactive oxygen species. *Scientific Reports*, 5:13961, 2015.

[272] M. F. Renschler. The emerging role of reactive oxygen species in cancer therapy. *European Journal of Cancer*, 40:1934–1940, 2004.

[273] S. C. Gupta, D. Hevia, S. Patchva, B. Park, W. Koh, and B. B. Aggarwal. Upsides and downsides of reactive oxygen species for cancer: The roles of reactive oxygen species in tumorigenesis, prevention, and therapy. *Antioxidants & Redox Signaling*, 16:1295–1322, 2012.

[274] D. B. Graves. The emerging role of reactive oxygen and nitrogen species in redox biology and some implications for plasma applications to medicine and biology. *Journal of Physics D: Applied Physics*, 45:263001, 2012.

[275] D. K. Böhme. Multiply-charged ions and interstellar chemistry. *Physical Chemistry Chemical Physics*, 13:18253–18263, 2011.

[276] S. Hammerum. Distonic radical cations in gaseous and condensed phase. *Mass Spectrometry Reviews*, 7:123–202, 1988.

[277] A. Verkhovtsev, E. Surdutovich, and A. V. Solov'yov. Multiscale approach predictions for biological outcomes in ion-beam cancer therapy. *Scientific Reports*, 6:27654, 2016.

[278] E. Surdutovich, A. V. Yakubovich, and A. V. Solov'yov. Biodamage via shock waves initiated by irradiation with ions. *Scientific Reports*, 3:1289, 2013.

[279] C. von Sonntag. *The Chemical Basis of Radiation Biology*. Taylor & Francis, New York, 1987.

[280] Farhataziz and M. A. J. Rodgers. *Radiation Chemistry: Principles and Applications*. VCH Publishers, New York, 1987.

[281] D. A. McCormick. GABA as an inhibitory neurotransmitter in human cerebral cortex. *Journal of Neurophysiology*, 62:1018–1027, 1989.

[282] M. Watanabe, K. Maemura, K. Oki, N. Shiraishi, Y. Shibayama, and K. Katsu. γ-aminobutyric acid (GABA) and cell proliferation: focus on cancer cells. *Histol. Hispatol.*, 21:1135–1141, 2006.

[283] S. Blanco, J. C. López, S. Mata, and J. L. Alonso. Conformations of γ-aminobutyric acid (GABA): the role of the n-π^* interaction. *Angewandte Chemie International Edition*, 49:9187–9192, 2010.

[284] L. Qiang, S. Cao, X. Zhao, X. Mao, Y. Guo, X. Liao, and Y. Zhao. Electrospray ionization tandem mass spectrometry differentiation of N-phosphoryl-α-, β-and γ-amino acids. *International Journal of Mass Spectrometry*, 266:42–49, 2007.

[285] L. M. Qiang, S. X. Cao, Y. C. Guo, X. C. Liao, and Y. F. Zhao. Interaction between N-phosphoryl, α, β -and γ-amino acids and nucleosides. *Chinese Chemical Letters*, 18:939 – 942, 2007.

[286] L. Qiang, R. Liu, Y. Xie, S. Cao, X. Liao, and Y. Zhao. A carbonyl oxygen migration of N-phosphoryl dipeptide methyl esters containing β-alanine or γ-amino butyric acid in electrospray ionization mass spectrometry. *International Journal of Mass Spectrometry*, 272:204–207, 2008.

[287] H. M. Ferraz, F. L. Pereira, E. R. Gonçalo, L. S. Santos, and M. N. Eberlin. Unexpected synthesis of conformationally restricted analogues of γ-amino butyric acid (GABA): mechanism elucidation by electrospray ionization mass spectrometry. *The Journal of organic chemistry*, 70:110–114, 2005.

[288] J. H. D. Eland. Dynamics of fragmentation reactions from peak shapes in multiparticle coincidence experiments. *Laser Chemistry*, 11:259–263, 1991.

[289] D. R. Dice and R. P. Steer. Evidence for the formation of unscavengeable diradical intermediates in the photolysis of cis- and trans-3-ethyl-2-propylthietan. *Journal of the Chemical Society, Chemical Communications*, pages 106–107, 1973.

[290] M. J. Polce and C. Wesdemiotis. The unimolecular chemistry of the 1,4-diradical $CH_2CH_2OCH_2$ in the gas phase. comparison to the distonic radical ions $CH_2CH_2OCH_2^+$ and $CH_2CH_2OCH_2$. *Journal of the American Chemical Society*, 115:10849–10856, 1993.

[291] R. Nádasdi, G. L. Zügner, M. Farkas, S. Dóbé, S. Maeda, and K. Morokuma. Photochemistry of methyl ethyl ketone: Quantum yields and s1/s0-diradical mechanism of photodissociation. *A European Journal of Chemical Physics and Physical Chemistry*, 11:3883–3895, 2010.

[292] E. Cadenas. *Oxidative Stress and Antioxidant Defenses in Biology*, chapter Mechanisms of Oxygen Activation and Reactive Oxygen Species Detoxification, pages 1–61. Springer US, Boston, MA, 1995.

[293] G. Poli, G. Leonarduzzi, F. Biasi, and E. Chiarpotto. Oxidative stress and cell signalling. *Current Medicinal Chemistry*, 11:1163–1182, 2004.

[294] E. Cadenas. Biochemistry of oxygen toxicity. *Annual Review of Biochemistry*, 58:79–110, 1989.

[295] V. Cadierno, M. Zablocka, B. Donnadieu, A. Igau, and J.-P. Majoral. Early transition metal diazoalkane complexes. *Angewandte Chemie International Edition*, 39:4524–4528, 2000.

[296] J. H. Jones. The mass spectra of amino-acid and peptide derivatives. *Quarterly Reviews of the Chemical Society*, 22:302–316, 1968.

[297] S. Denifl, I. Mähr, F. Ferreira da Silva, F. Zappa, T. D. Märk, and P. Scheier. Electron impact ionization studies with the amino acid valine in the gas phase and (hydrated) in helium droplets. *The European Physical Journal D*, 51:73–79, 2009.

[298] P. Papp, P. Shchukin, J. Kočíšek, and Š. Matejčík. Electron ionization and dissociation of aliphatic amino acids. *The Journal of Chemical Physics*, 137:105101, 2012.

[299] M. Tia, B. Cunha de Miranda, S. Daly, F. Gaie-Levrel, G. A. Garcia, L. Nahon, and I. Powis. Vuv photodynamics and chiral asymmetry in the photoionization of gas phase alanine enantiomers. *The Journal of Physical Chemistry A*, 118:2765–2779, 2014.

[300] O. Plekan, V. Feyer, R. Richter, M. Coreno, M. de Simone, K. C. Prince, and V. Carravetta. Investigation of the amino acids glycine, proline, and methionine by photoemission spectroscopy. *The Journal of Physical Chemistry A*, 111:10998–11005, 2007.

[301] A. Domaracka, M. Capron, S. Maclot, J.-Y. Chesnel, A. Méry, J.-C. Poully, J. Rangama, L. Adoui, P. Rousseau, and B. A. Huber. Ion interaction with biomolecular systems and the effect of the environment. *Journal of Physics: Conference Series*, 373:012005, 2012.

[302] V. Vizcaino, J.-C. Poully, J.-Y. Chesnel, A. Domaracka, S. Maclot, A. Méry, J. Rangama, P. Rousseau, L. Adoui, and B. Huber. Stability of multiply-charged biomolecular clusters formed upon interaction with low-energy highly charged ions. *International Journal of Mass Spectrometry*, 365:181–186, 2014.

[303] J. Kopyra, C. König-Lehmann, and E. Illenberger. Low energy electron attachment to N-acetylglycine. *Chemical Physics Letters*, 550:47 – 51, 2012.

[304] E. A. Syrstad and F. Turecek. Toward a general mechanism of electron capture dissociation. *Journal of the American Society for Mass Spectrometry*, 16:208 – 224, 2005.

[305] R. A. Zubarev, N. L. Kelleher, and F. W. McLafferty. Electron capture dissociation of multiply charged protein cations. a nonergodic process. *Journal of the American Chemical Society*, 120:3265–3266, 1998.

[306] J. Kopyra, C. König-Lehmann, and E. Illenberger. Electron attachment to the dipeptide alanyl-glycine. *Chemical Physics Letters*, 578:54 – 58, 2013.

[307] E. Alizadeh, D. Gschliesser, P. Bartl, M. Hager, A. Edtbauer, V. Vizcaino, A. Mauracher, M. Probst, T. D. Märk, S. Ptasińska, N. J. Mason, S. Denifl, and P. Scheier. Bond dissociation of the dipeptide dialanine and its derivative alanine anhydride induced by low energy electrons. *The Journal of Chemical Physics*, 134:054305, 2011.

[308] B. Puschnigg, S. E. Huber, M. Probst, K. Tanzer, V. Vizcaino, F. Ferreira da Silva, P. Scheier, P. Limao-Vieira, and S. Denifl. Electron attachment to the dipeptide dialanine: influence of methylation on site selective dissociation reactions. *Physical Chemistry Chemical Physics*, 15:3834–3840, 2013.

[309] K. Heyns and H.-F. Grützmacher. Massenspektrometrische untersuchungen, III. Massenspektren von freien und N-acetylierten aminosäuren. *Justus Liebigs Annalen der Chemie*, 667:194–204, 1963.

[310] R. Delaunay, J.-P. Champeaux, S. Maclot, M. Capron, A. Domaracka, A. Méry, B. Manil, L. Adoui, P. Rousseau, P. Moretto-Capelle, and B. Huber. Prompt and delayed fragmentation of bromouracil cations ionized by multiply charged ions. *The European Physical Journal D*, 68, 2014.

[311] C. Kurz, A. Mairani, and K. Parodi. First experimental-based characterization of oxygen ion beam depth dose distributions at the heidelberg ion-beam therapy center. *Physics in Medicine and Biology*, 57:5017, 2012.

[312] G. A. Jeffrey and W. Saenger. *Hydrogen Bonding in Biological Structures*. Springer-Verlag Berlin Heidelberg, 1 edition, 1991.

[313] G. R. Desiraju and T. Steiner. *The Weak Hydrogen Bond in Struct. Chemistry, Biology*, volume 123. Oxford University Press: Oxford and New York, 1999.

[314] H.-J. Schneider. Binding mechanisms in supramolecular complexes. *Angewandte Chemie International Edition*, 48:3924–3977, 2009.

[315] J. Tang, H. Yin, J. Qiu, M. J. Tucker, W. F. DeGrado, and F. Gai. Using two fluorescent probes to dissect the binding, insertion, and dimerization kinetics of a model membrane peptide. *Journal of the American Chemical Society*, 131:3816–3817, 2009.

[316] J. Laskin and J. H. Futrell. On the efficiency of energy transfer in collisional activation of small peptides. *The Journal of Chemical Physics*, 116:4302–4310, 2002.

[317] O. Abdelrahman, B. Grotemeyer, and J. Grotemeyer. The mobile proton in biomolecular clusters: Tripeptides and vanillic acid. *International Journal of Mass Spectrometry*, 354-355:398–405, 2013.

[318] I. Alkorta, J. Elguero, S. Goswami, and R. Mukherjee. Interaction of adenine with synthetic receptors: a theoretical study. *Journal of the Chemical Society, Perkin Transactions 2*, 5:894–898, 2002.

[319] I. I. Schuster. Stereoelectronic substituent effects on intramolecular h-bonding in crowded o-anisic acids. *Journal of the Chemical Society, Perkin Transactions 2*, pages 1961–1966, 2002.

[320] B. Over, P. McCarren, P. Artursson, M. Foley, F. Giordanetto, G. Grönberg, C. Hilgendorf, I. Maurice D. Lee, P. Matsson, G. Muncipinto, M. Pellisson, M. W. D. Perry, R. Svensson, J. R. Duvall, and J. Kihlberg. Impact of stereospecific intramolecular hydrogen bonding on cell permeability and physicochemical properties. *Journal of Medicinal Chemistry*, 57:2746–2754, 2014.

[321] O. Khakshoor and J. S. Nowick. Artificial β-sheets: chemical models of β-sheets. *Current opinion in chemical biology*, 12:722–729, 2008.

[322] S. T. Phillips, G. Piersanti, and P. A. Bartlett. Quantifying amino acid conformational preferences and side-chain-side-chain interactions in β-hairpins. *Proceedings of the National Academy of Sciences of the United States of America*, 102:13737–13742, 2005.

[323] M. C. Hammond, B. Z. Harris, W. A. Lim, and P. A. Bartlett. β strand peptidomimetics as potent {PDZ} domain ligands. *Chemistry and Biology*, 13:1247–1251, 2006.

[324] O. Khakshoor, B. Demeler, and J. S. Nowick. Macrocyclic β-sheet peptides that mimic protein quaternary structure through intermolecular β-sheet interactions. *Journal of the American Chemical Society*, 129:5558–5569, 2007.

[325] R. J. Plowright, E. Gloaguen, and M. Mons. Compact folding of isolated four-residue neutral peptide chains: H-bonding patterns and entropy effects. *A European Journal of Chemical Physics and Physical Chemistry*, 12:1889–1899, 2011.

[326] S. Jaeqx, J. Oomens, A. Cimas, M.-P. Gaigeot, and A. M. Rijs. Gas-phase peptide structures unraveled by far-IR spectroscopy: Combining IR-UV ion-dip experiments with Born-Oppenheimer molecular dynamics simulations. *Angewandte Chemie International Edition*, 53:3663–3666, 2014.

[327] M. Raszka and N. O. Kaplan. Intramolecular hydrogen bonding in flavin adenine dinucleotide. *Proceedings of the National Academy of Sciences*, 71:4546–4550, 1974.

[328] J. S. Nowick. Exploring β-sheet structure and interactions with chemical model systems. *Accounts of Chemical Research*, 41:1319–1330, 2008.

[329] N. E. Tayar, A. E. Mark, P. Vallat, R. M. Brunne, B. Testa, and W. F. van Gunsteren. Solvent-dependent conformation and hydrogen-bonding capacity of cyclosporin a: evidence from partition coefficients and molecular dynamics simulations. *Journal of Medicinal Chemistry*, 36:3757–3764, 1993.

[330] T. Rezai, J. E. Bock, M. V. Zhou, C. Kalyanaraman, R. S. Lokey, , and M. P. Jacobson. Conformational flexibility, internal hydrogen bonding, and passive membrane permeability: Successful in silico prediction of the relative permeabilities of cyclic peptides. *Journal of the American Chemical Society*, 128:14073–14080, 2006.

[331] D. G. Fleming, J. Manz, K. Sato, and T. Takayanagi. Fundamental change in the nature of chemical bonding by isotopic substitution. *Angewandte Chemie International Edition*, 53:13706–13709, 2014.

[332] C. Pérez, M. T. Muckle, D. P. Zaleski, N. A. Seifert, B. Temelso, G. C. Shields, Z. Kisiel, and B. H. Pate. Structures of cage, prism, and book isomers of water hexamer from broadband rotational spectroscopy. *Science*, 336:897–901, 2012.

[333] L. De Marco, M. Thämer, M. Reppert, and A. Tokmakoff. Direct observation of intermolecular interactions mediated by hydrogen bonding. *The Journal of Chemical Physics*, 141:034502, 2014.

[334] R. Jimenez, G. R. Fleming, P. V. Kumar, and M. Maroncelli. Femtosecond solvation dynamics of water. *Nature*, 369:471–473, 1994.

[335] G. R. Fleming and M. Cho. Chromophore-solvent dynamics. *Annual Review of Physical Chemistry*, 47:109–134, 1996.

[336] S. Park, J. Kim, and N. F. Scherer. Two-dimensional measurements of the solvent structural relaxation dynamics in dipolar solvation. *Physical Chemistry Chemical Physics*, 14:8116–8122, 2012.

[337] G. P. Dado and S. H. Gellman. Intramolecular hydrogen bonding in derivatives of β-alanine and γ-amino butyric acid; model studies for the folding of unnatural polypeptide backbones. *Journal of the American Chemical Society*, 116:1054–1062, 1994.

[338] C. A. Schalley and P. Weis. Unusually stable magic number clusters of serine with a surprising preference for homochirality. *International Journal of Mass Spectrometry*, 221:9–19, 2002.

[339] B. Kuhn, P. Mohr, and M. Stahl. Intramolecular hydrogen bonding in medicinal chemistry. *Journal of Medicinal Chemistry*, 53:2601–2611, 2010.

[340] Y. Zhao and D. G. Truhlar. Design of density functionals that are broadly accurate for thermochemistry, thermochemical kinetics, and nonbonded interactions. *The Journal of Physical Chemistry A*, 109:5656–5667, 2005.

[341] S. Grimme. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *Journal of Computational Chemistry*, 27:1787–1799, 2006.

[342] M. Modrzejewski, M. Lesiuk, L. Rajchel, M. M. Szczęśniak, and G. Chałasiński. A first-principles-based correlation functional for harmonious connection of short-range correlation and long-range dispersion. *The Journal of Chemical Physics*, 137:204121, 2012.

[343] A. E. Reed, L. A. Curtiss, and F. Weinhold. Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint. *Chemical Reviews*, 88:899–926, 1988.

[344] E. R. Johnson, S. Keinan, P. Mori-Sánchez, J. Contreras-García, A. J. Cohen, and W. Yang. Revealing noncovalent interactions. *Journal of the American Chemical Society*, 132:6498–6506, 2010.

[345] J. Contreras-García, E. R. Johnson, S. Keinan, R. Chaudret, J.-P. Piquemal, D. N. Beratan, and W. Yang. NCIPLOT: A program for plotting noncovalent interaction regions. *Journal of Chemical Theory and Computation*, 7:625–632, 2011.

[346] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, 81:3684, 1984.

[347] R. J. Loncharich, B. R. Brooks, and R. W. Pastor. Langevin dynamics of peptides: The frictional dependence of isomerization rates of N-acetylalanyl-N-methylamide. *Biopolymers*, 32:523–535, 1992.

[348] B. Keller, Z. Gattin, and W. F. van Gunsteren. What stabilizes the 3_{14}-helix in β^3-peptides? a conformational analysis using molecular simulation. *Proteins*, 78:1677–1690, 2010.

[349] T. A. Halgren. Merck molecular force field. i. basis, form, scope, parameterization, and performance of mmff94. *Journal of Computational Chemistry*, 17:490–519, 1996.

[350] W. Hehre, P. Klunzinger, B. Deppmeier, A. Driessen, and J. Johnson. Spartan14, 2014.

[351] Y. Shao, L. F. Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S. T. Brown, A. T. Gilbert, L. V. Slipchenko, S. V. Levchenko, D. P. O'Neill, R. A. DiStasio Jr, R. C. Lochan, T. Wang, G. J. Beran, N. A. Besley, J. M. Herbert, C. Yeh Lin, T. Van Voorhis, S. Hung Chien, A. Sodt, R. P. Steele, V. A. Rassolov, P. E. Maslen, P. P. Korambath, R. D. Adamson, B. Austin, J. Baker, E. F. C. Byrd, H. Dachsel, R. J. Doerksen, A. Dreuw, B. D.

Dunietz, A. D. Dutoi, T. R. Furlani, S. R. Gwaltney, A. Heyden, S. Hirata, C.-P. Hsu, G. Kedziora, R. Z. Khalliulin, P. Klunzinger, A. M. Lee, M. S. Lee, W. Liang, I. Lotan, N. Nair, B. Peters, E. I. Proynov, P. A. Pieniazek, Y. Min Rhee, J. Ritchie, E. Rosta, C. David Sherrill, A. C. Simmonett, J. E. Subotnik, H. Lee Woodcock III, W. Zhang, A. T. Bell, A. K. Chakraborty, D. M. Chipman, F. J. Keil, A. Warshel, W. J. Hehre, H. F. Schaefer III, J. Kong, A. I. Krylov, P. M. W. Gill, and M. Head-Gordon. Advances in methods and algorithms in a modern quantum chemistry program package. *Physical Chemistry Chemical Physics*, 8:3172–3191, 2006.

[352] E. D. Glendening, A. E. Reed, J. E. Carpenter, and F. Weinhold. NBO version 3.1.

[353] F. B. van Duijneveldt, J. G. C. M. van Duijneveldt-van de Rijdt, and J. H. van Lenthe. State of the art in counterpoise theory. *Chemical Reviews*, 94:1873–1885, 1994.

[354] L. A. Curtiss, P. C. Redfern, K. Raghavachari, and J. A. Pople. Gaussian-3x (g3x) theory: Use of improved geometries, zero-point energies, and Hartree-Fock basis sets. *The Journal of Chemical Physics*, 114:108–117, 2001.

[355] G. A. Jeffrey. *An Introduction to Hydrogen Bonding (Topics in Physical Chemistry)*. Oxford University Press, 1997.

[356] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *Journal of Physical Chemistry*, 98:11623–11627, 1994.

[357] Y. Zhao and D. G. Truhlar. Hybrid meta density functional theory methods for thermochemistry, thermochemical kinetics, and noncovalent interactions: The MPW1B95 and MPWB1K models and comparative assessments for hydrogen bonding and van der waals interactions. *The Journal of Physical Chemistry A*, 108:6908–6918, 2004.

[358] E. Gloaguen, B. de Courcy, J.-P. Piquemal, J. Pilmé, O. Parisel, R. Pollet, H. S. Biswal, F. Piuzzi, B. Tardivel, M. Broquier, and M. Mons. Gas-phase folding of a two-residue model peptide chain: On the importance of an interplay between experiment and theory. *Journal of the American Chemical Society*, 132:11860–11863, 2010.

[359] S. Simon, M. Duran, and J. J. Dannenberg. How does basis set superposition error change the potential surfaces for hydrogen-bonded dimers? *The Journal of Chemical Physics*, 105:11024–11031, 1996.

[360] P. Salvador, B. Paizs, M. Duran, and S. Suhai. On the effect of the bsse on intermolecular potential energy surfaces. comparison of a priori and a

posteriori bsse correction schemes. *Journal of Computational Chemistry*, 22:765–786, 2001.

[361] S. H. Gellman, G. P. Dado, G. B. Liang, and B. R. Adams. Conformation-directing effects of a single intramolecular amide-amide hydrogen bond: variable-temperature NMR and IR studies on a homologous diamide series. *Journal of the American Chemical Society*, 113:1164–1173, 1991.

[362] E. Espinosa, E. Molins, and C. Lecomte. Hydrogen bond strengths revealed by topological analyses of experimentally observed electron densities. *Chemical Physics Letters*, 285:170–173, 1998.

[363] G. Sánchez-Sanz, C. Trujillo, I. Alkorta, and J. Elguero. Weak interactions between hypohalous acids and dimethylchalcogens. *Physical Chemistry Chemical Physics*, 14:9880–9889, 2012.

[364] G. Sánchez-Sanz, C. Trujillo, I. Alkorta, and J. Elguero. Intermolecular weak interactions in htexh dimers (x=o, s, se, te): Hydrogen bonds, chalcogen-chalcogen contacts and chiral discrimination. *A European Journal of Chemical Physics and Physical Chemistry*, 13:496–503, 2012.

[365] G. Sánchez-Sanz, C. Trujillo, and I. Alkorta. Structure, binding energy and chiral discrimination in oxathiirane homodimers. *Computational and Theoretical Chemistry*, 1090:171–179, 2016.

[366] K. Wendler, J. Thar, S. Zahn, and B. Kirchner. Estimating the hydrogen bond energy. *The Journal of Physical Chemistry A*, 114:9529–9536, 2010.

[367] P. C. Hariharan and J. A. Pople. The influence of polarization functions on molecular orbital hydrogenation energies. *Theoretica chimica acta*, 28:213–222, 1973.

[368] M. Alcamí, O. Mó, and M. Yáñez. G2 ab initio calculations on three-membered rings: Role of hydrogen atoms. *Journal of Computational Chemistry*, 19:1072–1086, 1998.

[369] A. E. Aliev, E. J. MacLean, K. D. M. Harris, B. M. Kariuki, and C. Glidewell. Dynamics of the hydrogen-bonding arrangement in solid triphenylmethanol: An investigation by solid-state 2H NMR spectroscopy. *The Journal of Physical Chemistry B*, 102:2165–2175, 1998.

[370] D. J. Plaut, K. T. Holman, A. M. Pivovar, and M. D. Ward. Building molecular frameworks with tailored pore structures. *Journal of Physical Organic Chemistry*, 13:858–869, 2000.

[371] M. A. Williams, J. M. Goodfellow, and J. M. Thornton. Buried waters and internal cavities in monomeric proteins. *Protein Science*, 3:1224–1235, 1994.

[372] I. Alkorta, J. Elguero, M. Yáñez, and O. Mó. Cooperativity in beryllium bonds. *Physical Chemistry Chemical Physics*, 16:4305–4312, 2014.

[373] L. P. Wolters, N. W. G. Smits, and C. F. Guerra. Covalency in resonance-assisted halogen bonds demonstrated with cooperativity in N-halo-guanine quartets. *Physical Chemistry Chemical Physics*, 17:1585–1592, 2015.

[374] L. Albrecht, R. J. Boyd, O. Mo, and M. Yanez. Cooperativity between hydrogen bonds and beryllium bonds in (H2O)nBeX_2 (n = 1-3, x = H,F) complexes. a new perspective. *Physical Chemistry Chemical Physics*, 14:14540–14547, 2012.

[375] L. Albrecht and R. J. Boyd. Atomic energy analysis of cooperativity, anti-cooperativity, and non-cooperativity in small clusters of methanol, water, and formaldehyde. *Computational and Theoretical Chemistry*, 1053:328 – 336, 2015.

[376] I. Alkorta, J. Elguero, O. Mó, M. Yáñez, and J. E. Del Bene. Using beryllium bonds to change halogen bonds from traditional to chlorine-shared to ion-pair bonds. *Physical Chemistry Chemical Physics*, 17:2259–2267, 2015.

[377] H. Kleeberg, D. Klein, and W. A. P. Luck. Quantitative infrared-spectroscopic investigations of hydrogen-bond cooperativity. *The Journal of Physical Chemistry*, 91:3200–3203, 1987.

[378] N. Juranić, V. A. Likić, F. G. Prendergast, and S. Macura. Protein-solvent hydrogen bonding studied by NMR 1JNC' coupling constant determination and molecular dynamics simulations. *Journal of the American Chemical Society*, 118:7859–7860, 1996.

[379] J. S. Miller, R. J. Kennedy, and D. S. Kemp. Solubilized, spaced poly-alanines: A context-free system for determining amino acid α-helix propensities. *Journal of the American Chemical Society*, 124:945–962, 2002.

[380] N. A. Seifert, D. P. Zaleski, C. Pérez, J. L. Neill, B. H. Pate, M. Vallejo-López, A. Lesarri, E. J. Cocinero, F. Castaño, and I. Kleiner. Probing the c-h—π weak hydrogen bond in anesthetic binding: The sevoflurane-benzene cluster. *Angewandte Chemie International Edition*, 53:3210–3213, 2014.

[381] R. Wieczorek and J. J. Dannenberg. Hydrogen-bond cooperativity, vibrational coupling, and dependence of helix stability on changes in amino acid sequence in small 310-helical peptides. a density functional theory study. *Journal of the American Chemical Society*, 125:14065–14071, 2003.

[382] R. Wieczorek and J. J. Dannenberg. Enthalpies of hydrogen-bonds in α-helical peptides. an ONIOM DFT/AM1 study. *Journal of the American Chemical Society*, 127:14534–14535, 2005.

[383] R. Wieczorek and J. J. Dannenberg. Comparison of fully optimized α- and 310-helices with extended β-strands. an oniom density functional theory study. *Journal of the American Chemical Society*, 126:14198–14205, 2004.

[384] J. Li, Y. Wang, J. Chen, Z. Liu, A. Bax, and L. Yao. Observation of α-helical hydrogen-bond cooperativity in an intact protein. *Journal of the American Chemical Society*, 138:1824–1827, 2016.

[385] R. Viswanathan, A. Asensio, and J. J. Dannenberg. Cooperative hydrogen-bonding in models of antiparallel β-sheets. *The Journal of Physical Chemistry A*, 108:9205–9212, 2004.

[386] G. Sánchez-Sanz, C. Trujillo, I. Alkorta, and J. Elguero. Competition between intramolecular hydrogen and pnictogen bonds in protonated systems. *Theoretical Chemistry Accounts*, 135:140, 2016.

[387] F. Weinhold. Nature of H-bonding in clusters, liquids, and enzymes: an ab initio, natural bond orbital perspective. *Journal of Molecular Structure: THEOCHEM*, 398-399:181–197, 1997.

[388] A. S. Mahadevi and G. N. Sastry. Cooperativity in noncovalent interactions. *Chemical Reviews*, 116:2775–2825, 2016.

[389] J. C. Tully. Molecular dynamics with electronic transitions. *The Journal of Chemical Physics*, 93:1061–1071, 1990.

[390] L. Zhao and C.-J. Li. Functionalizing glycine derivatives by direct CC bond formation. *Angewandte Chemie*, 120, 2008.

[391] P. Linstrom and W. Mallard. *NIST Chemistry WebBook,* http://webbook.nist.gov/, *(retrieved February 29, 2016)*, volume l. NIST Standard Reference Database Number 69, National Institute of Standards and Technology, Gaithersburg MD, 20899,, 1990.

[392] M. A. Zwijnenburg, G. Cheng, T. O. McDonald, K. E. Jelfs, J.-X. Jiang, S. Ren, T. Hasell, F. Blanc, A. I. Cooper, and D. J. Adams. Shedding light on structure-property relationships for conjugated microporous polymers: The importance of rings and strain. *Macromolecules*, 46:7696–7704, 2013.

[393] W. Zhang, T. Hou, C. S. nad Wilson S. Ross, and D. A. Case. Leap and gleap manual for ambertools, 2008.

[394] H. C. Andersen. Molecular dynamics simulations at constant pressure and/or temperature. *The Journal of Chemical Physics*, 72:2384–2393, 1980.

Part V

Appendices

Appendices

Additional information for thymidine$^+$

A.1 Experimental details

The experiments have been performed at ARIBE, the low-energy ions facility of GANIL (Caen, France) and at the GASPHASE beamline of the synchrotron radiation facility ELETTRA (Trieste, Italy). Both experiments are based on crossed-beam set-ups using coincidence time-of-flight mass spectrometry. The effusive beam of neutral thymidine molecule (2'-deoxy-thymidine) was produced by heating a powder (Sigma-Aldrich - 99% purity) in a molybdenum oven at 130°C. At this temperature, the thermal decomposition is avoided.

Ion collision experiment

In the ion collision experiment, the ion beam of O^{6+} at 48 keV is produced by an ECR source then pulsed in bunches of 500 ns and transported to the interaction region. The charged products resulting from the interaction of thymidine molecules with the ion beam are analysed using a Wiley-McLaren linear time-of-flight mass spectrometer coupled with a Daly-type detector allowing a constant detection efficiency over a large mass range. The set-up allows also to measure fragment ion-fragment ion coincidences in order to study the fragmentation dynamics as a function of the initial charge state of the molecule.
Some peaks in the mass spectrum, shown in the Fig.1(b), are due to a contamination of the residual gas by previous experiments on PAH molecules (namely pyrene with mass 202 and anthracene with mass 178).

Photoionisation experiment

The end-station used in the photoionisation experiment consists of a Wiley-McLaren type time-of-flight mass spectrometer coupled to a 150 mm mean radius electron energy analyser (VG 220i, with six electron multiplier detectors) in order to perform photoelectron spectroscopy (PES), photoion mass spectrometry (PIMS) and

photoelectron-photoion coincidence measurements (PEPICO). In PES experiment the pass energy of the VG hemispherical analyser was set at 5 eV corresponding to an energy resolution FWHM of 200 meV, while in the PEPICO it was 20 eV corresponding to FWHM of 800 meV.

A.2 Fitting Method

The overall mass spectra obtained by different methods look similar, the same peaks are observed differing only by their relative intensity. The range of binding energies E_b covered by the PEPICO experiments allowed to sample molecular orbitals up to $E_b = 15.74$ eV, therefore not all of the fragments observed in the ion spectrum (Fig.6.1b) are also observed in the photoionisation measurements (Fig.1h-k). In particular, more energetic processes induced by inner valence ionisation and resulting in multiple H-losses, formation of smaller fragments (m/z $<$ 50 amu) and fragments with a broad kinetic energy distribution have a reduced contribution or are completely absent in the PEPICO spectra. This corresponds to the tail towards larger excitation energy in Fig.6.2. Thus further studies should include PEPICO spectra extending to higher excitation energies to fully represent the ion-induced mass spectrum.
In the present we have selected the 11 most intense peaks, present in both sets of data, to be implemented in the fitting procedure (Tab.A.1).
The areas of the peaks of these eleven fragments in the ion (I_k) and in the thirteen PEPICO (P_{jk}) spectra have been evaluated by fitting Gaussian functions to their respective mass spectra. Then via a constrained linear least squares fitting procedure:

$$I_k = \sum_{j=1}^{13} A_j \cdot P_{ik} \tag{A.1}$$

where $k = 1 - 11$ is the index of the eleven charged species from m/z = 31 to 242 amu (see Tab.A.1 for the list) and $j = 1 - 13$ is the index of the thirteen measured PEPICO spectra, corresponding to increasing binding energies from 7.96 to 15.74 eV (see the inset label of Fig.A.2 for the full list), we searched for the best values of the A_j factors that reproduce the intensity I_k of the ion fragments in the ion impact spectrum.
In the fitting procedure, we *i)* constrained $A_j \geq 0$ and *ii)* initialised the fit with a vector of thirteen random numbers. In order to sample a broad range of the multidimensional surface of the parameters, an arbitrary number of 10,000 repetitions of the fitting procedure, each one with a different set of initial parameters, has been performed. The result is reported in Fig.A.1, together with the weight averaged χ^2 of all these solutions.
The frequency distribution of each one of the parameters A_j is displayed in Fig.A.2 versus its value, demonstrating the good convergence of the fitting procedure. The

Table A.1: List of charged species implemented in the fitting procedure.

m/z	Assignment
31	$HOCH_2^+$
55	$C_3H_3O^+$
73	$C_3H_5O_2^+$
98	$C_3H_2N_2O_2^+$ $C_4H_2O_3^+$ $C_5H_6O_2^+$
99	$C_3H_3N_2O_2^+$ $C_4H_3O_3^+$ $C_5H_7O_2^+$
110	$C_5H_6N_2O^+$
116	$(\mathbf{S\text{-}1})^+$
117	$\mathbf{S}^+$
126	$(\mathbf{B}+\mathbf{1})^+$
127	$(\mathbf{B}+\mathbf{2})^+$
242	$dThy^+$

half width of each distribution provides the uncertainty of the parameters, as reported in Fig.6.2 and Fig.A.1

The comparison between the measured (I_k) and "reconstructed" (I_k(fit)) distribution of intensity of the selected charged species in the ion spectrum is reported in Fig.A.3. I_k(fit) is calculated according to Eq.A.1 and using the values of the A_j parameters corresponding to the black crosses in Fig.A.1.

A.3 Molecular Orbitals

We show in Fig.A.4 the highest occupied molecular orbitals (HOMOs) of thymidine computed with the Hartree-Fock method and with a basis set 6-311G(d,p). The energies in the figure corresponds to the ionisation values computed with the outer-valance Green's function (OVGF) method.

A.4 Potential energy surfaces

The extensive potential energy surface exploration is presented in Figs.A.5-A.14. The energies in these figures are given in eV in respect to neutral thymidine, calculated at DFT-B3LYP/6-311G(d,p) level of theory. The most relevant pathways corresponds to the main peaks observed experimentally and used as well for the fitting procedure.

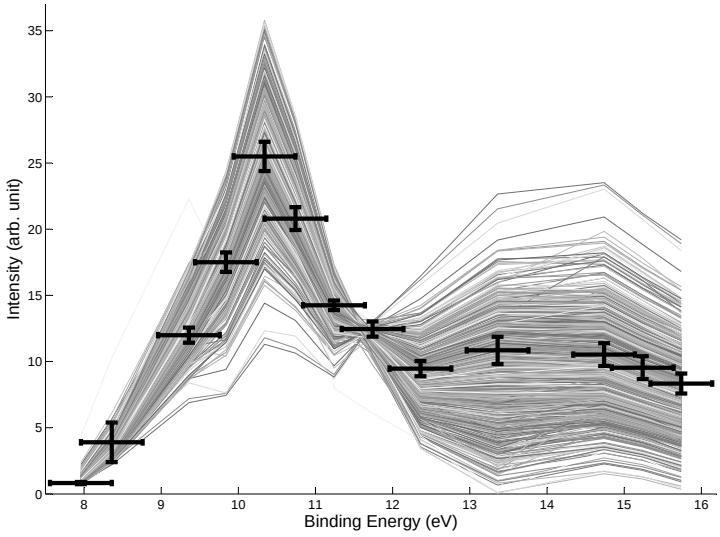


Figure A.1: The grey lines show the best values of the A_j parameters in the 10,000 repetitions of the fitting procedure reported versus binding energy. The black crosses are the weight averaged χ^2 of these solutions, reported with uncertainties in both binding energy (due to energy resolution in the PEPICO experiments) and averaged value (see text).

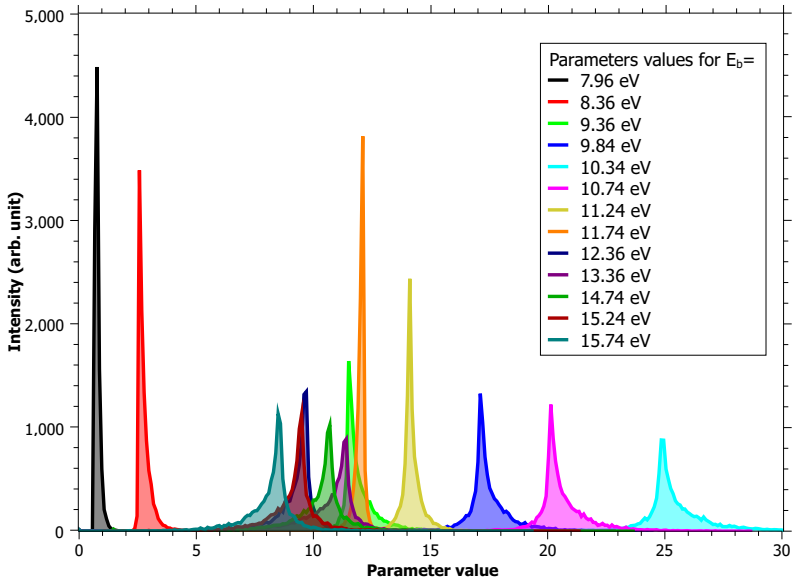


Figure A.2: The frequency distribution of the 13 A_j parameters reported versus their value. The correspondence between each parameter and the binding energy of the PEPICO spectra is given in the panel. The area of each one of these peaks is equal to 10,000 (number of fitting repetitions) while the half width provides an estimate of the uncertainty.

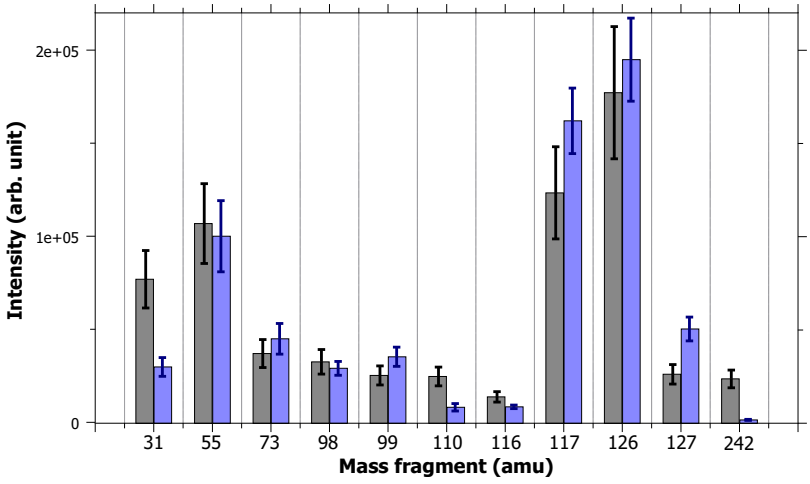


Figure A.3: Comparison between the measured intensity of the selected charged species in the ion spectrum (I_k, grey bars) and the one reconstructed by eq.A.1 (I_k(fit), blue bars), using the the best values of the A_j parameters reported in Fig.6.2 and Fig.A.1

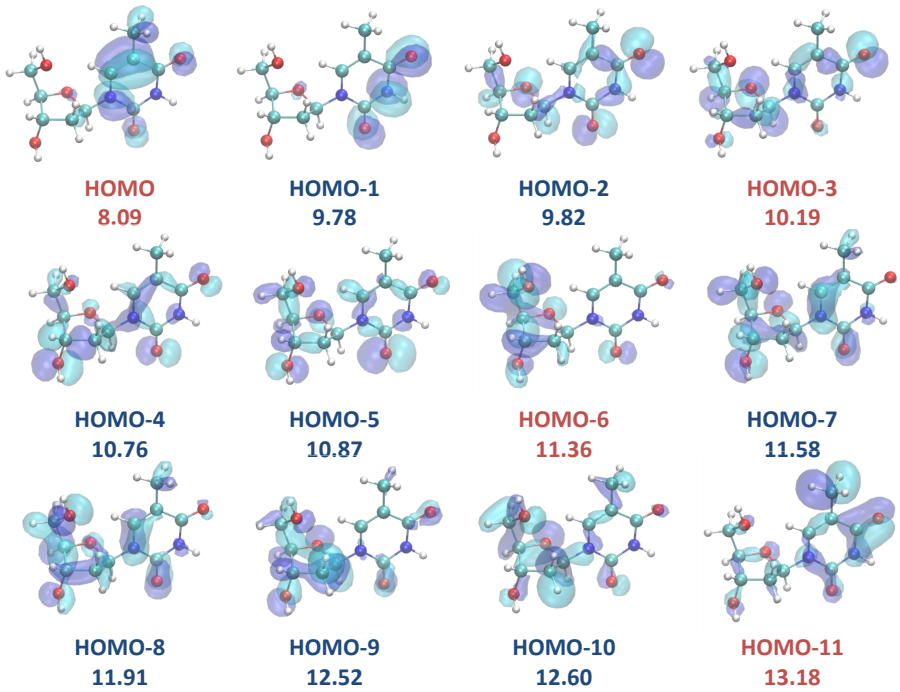


Figure A.4: Computed electron densities of the 12 highest occupied molecular orbitals (HOMOs). Ionisation energies from each HOMO are given in eV. In red the orbitals discussed in the Chap.6.

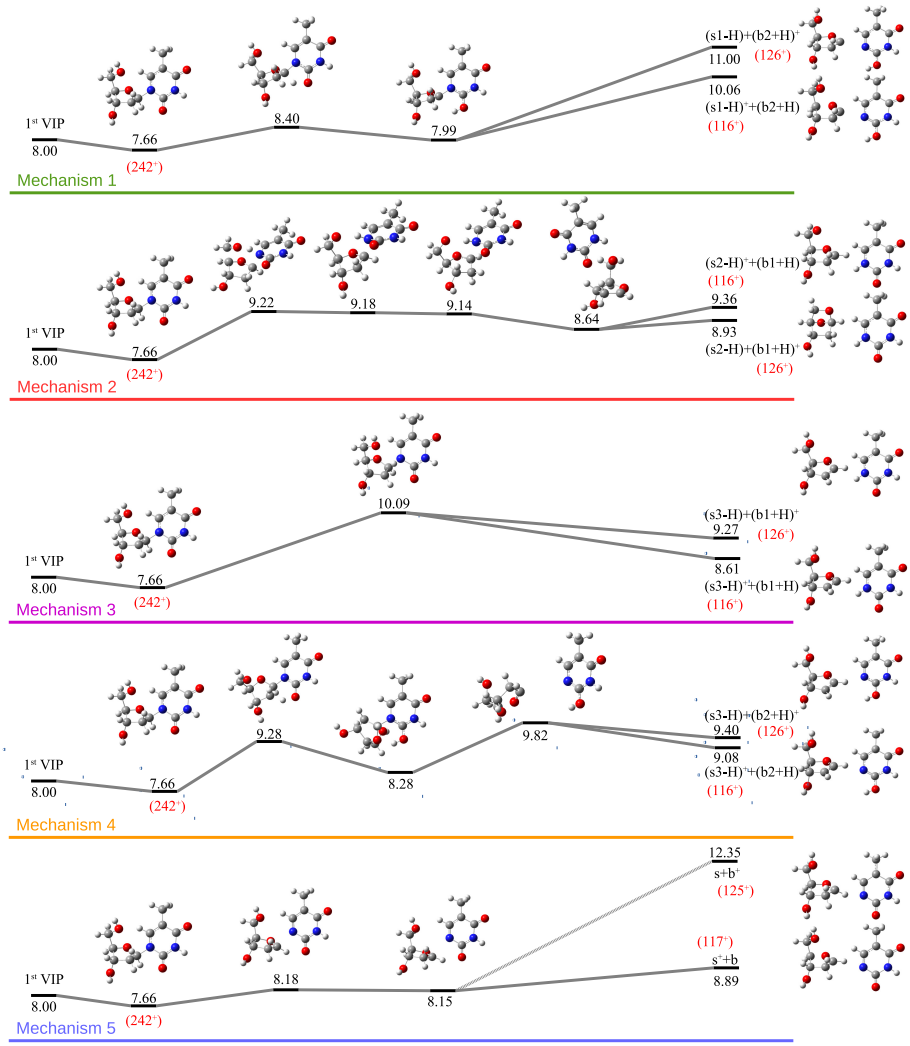


Figure A.5: Fragmentation pathways of singly charge thymidine. Relative energies, δE, are given in eV at the DFT-B3LYP/6-311G(d,p) level of theory including the zero point energy correction with respect to the neutral thymidine.

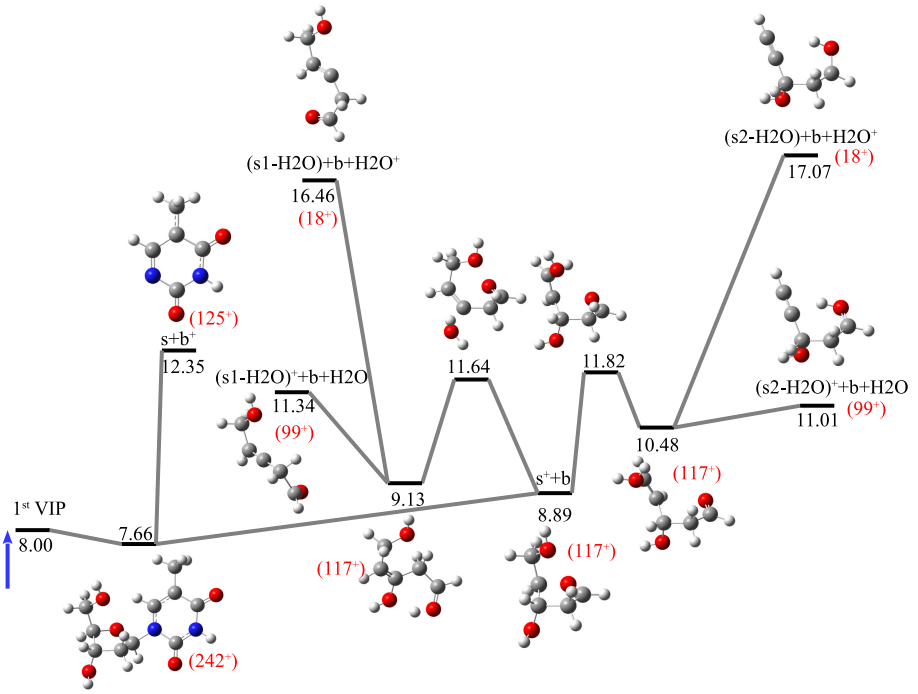


Figure A.6: See the caption of Fig.A.5

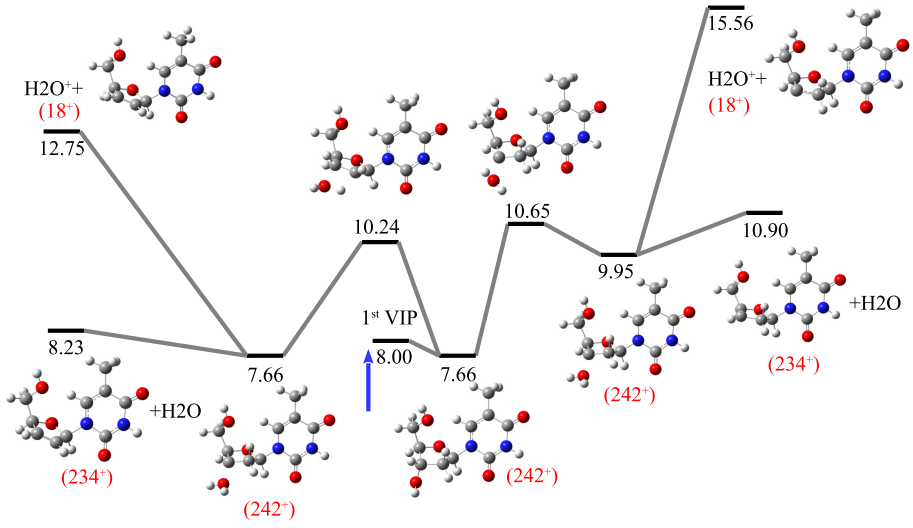


Figure A.7: See the caption of Fig.A.5

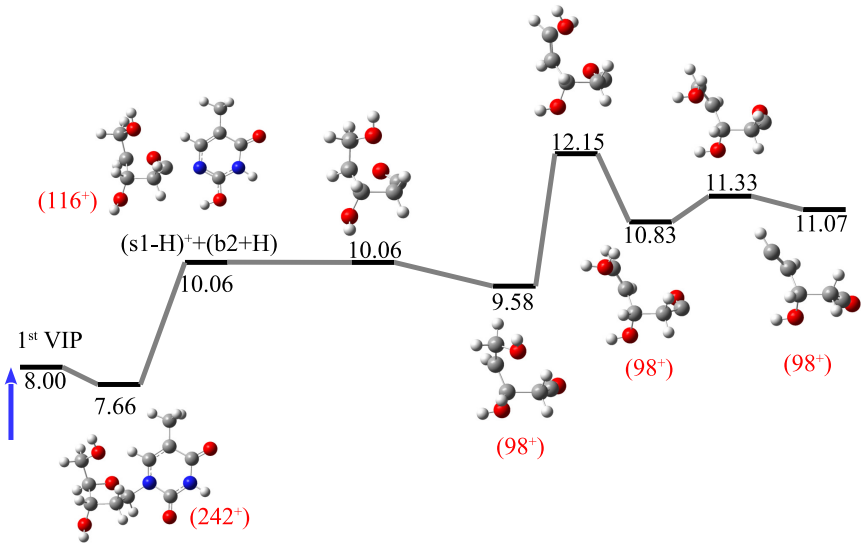


Figure A.8: See the caption of Fig.A.5

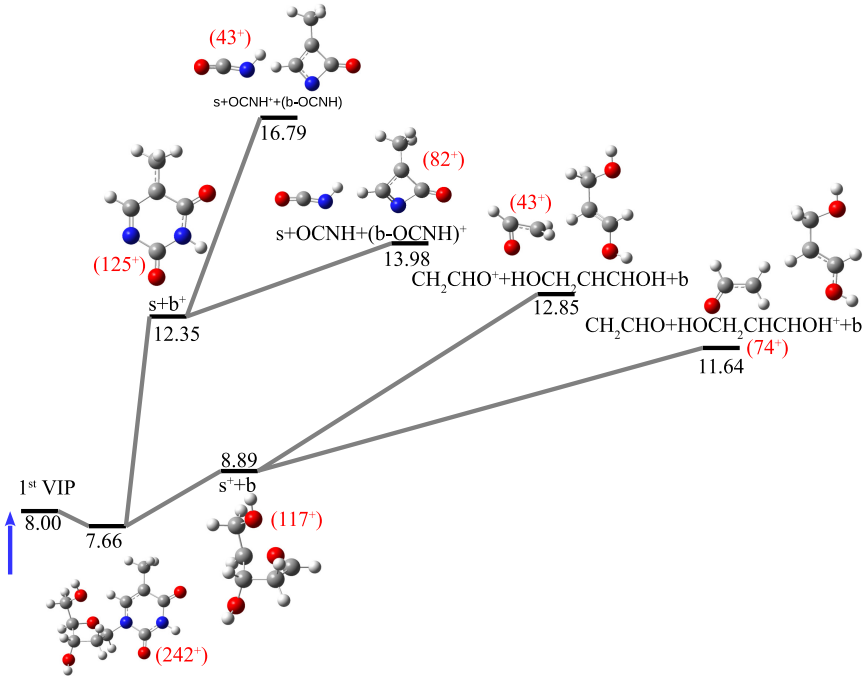


Figure A.9: See the caption of Fig.A.5

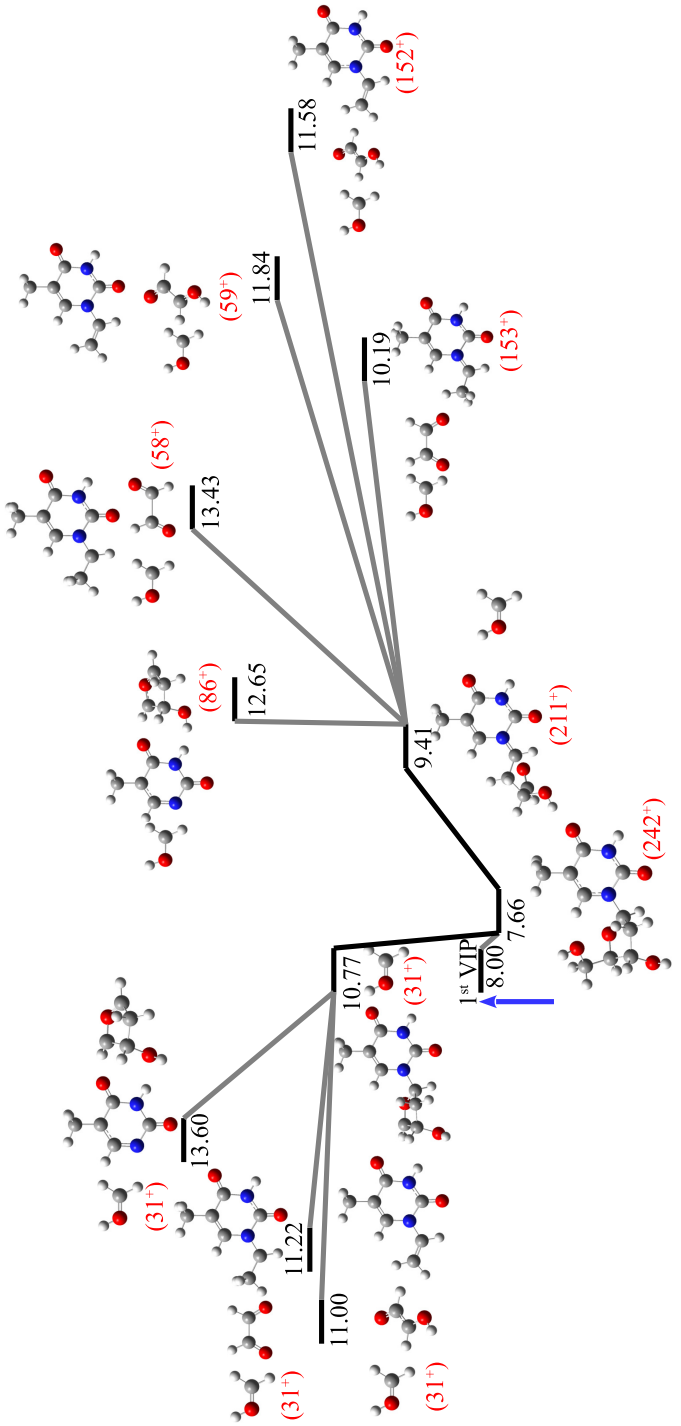


Figure A.10: See the caption of Fig.A.5

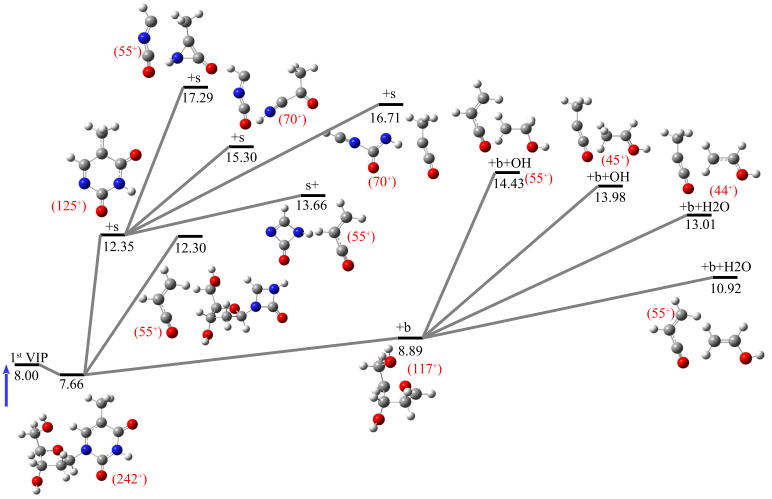


Figure A.11: See the caption of Fig.A.5

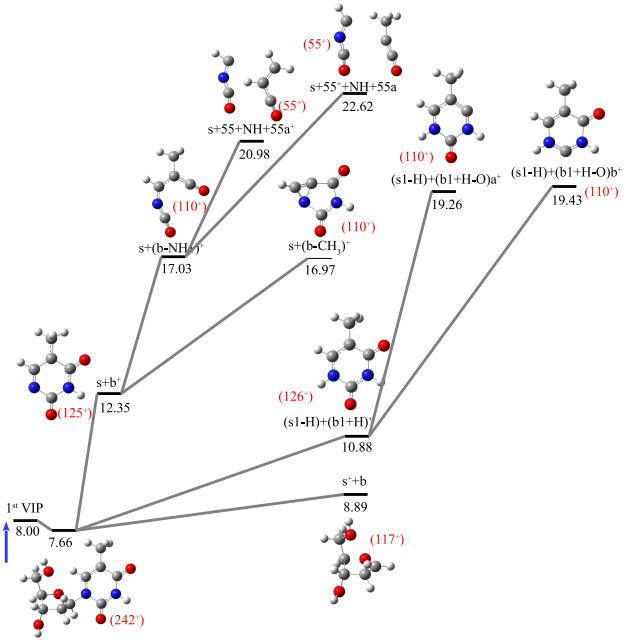


Figure A.12: See the caption of Fig.A.5

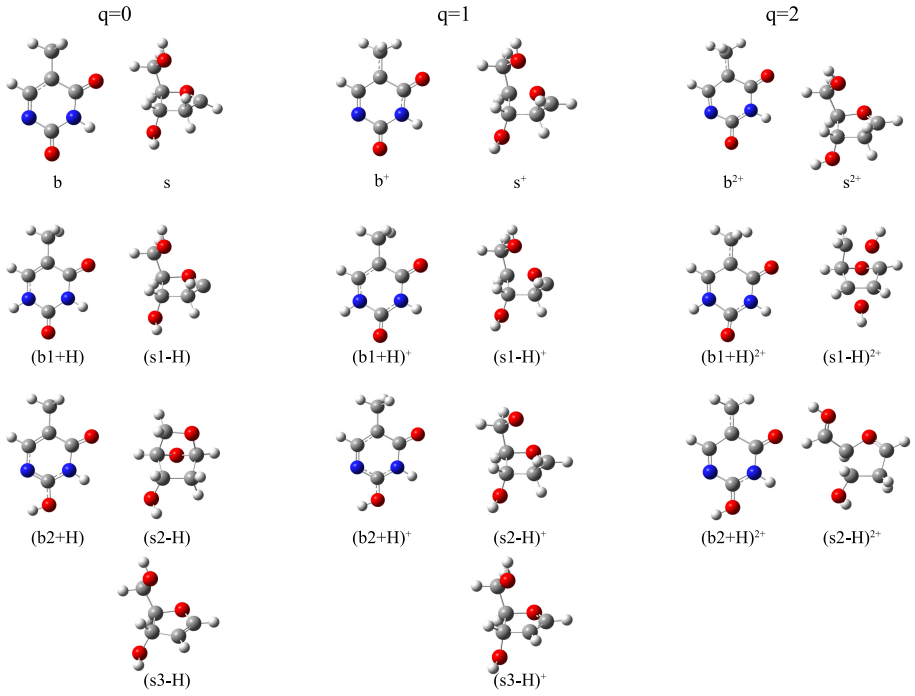


Figure A.13: The main products among the heaviest fragments of neutral, singly and doubly charged sugar part (s), base part (b) and these two with and without hydrogen atom optimised at DFT-B3LYP/6-311G(d,p) level of theory.

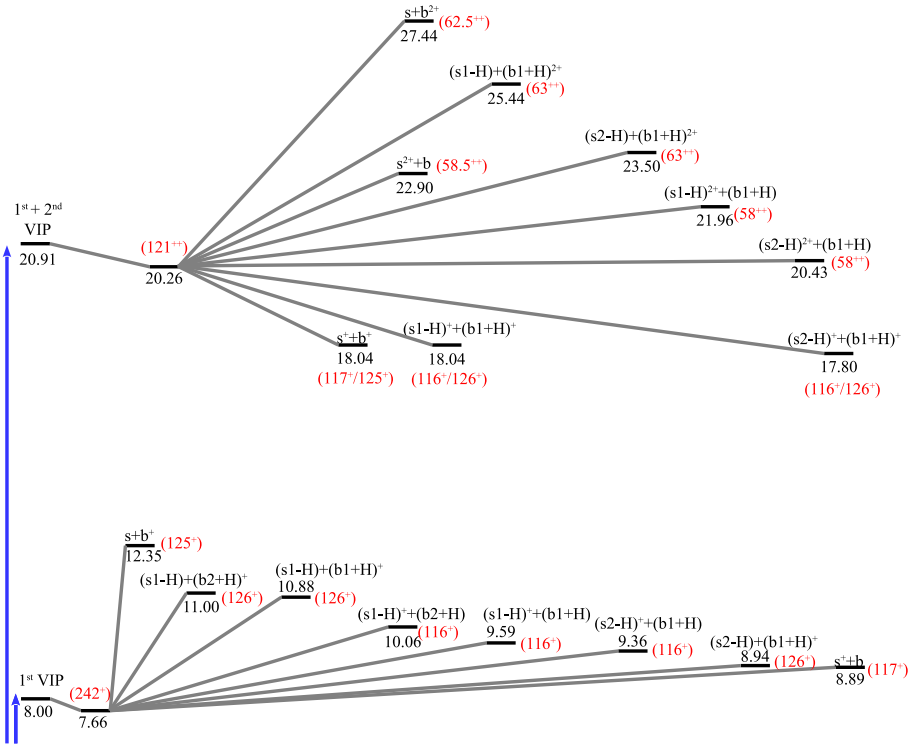
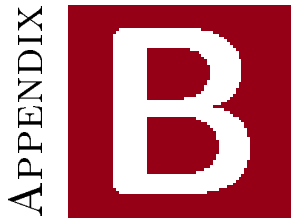


Figure A.14: Energy levels of the fragmentation pathways for singly and doubly charged thymidine leading to the main products among the heaviest singly and doubly charged fragments calculated at DFT-B3LYP/6-311G(d,p) level of theory. Relative energies in eV in respect to the neutral thymidine molecule.



Appendices
Additional information for glycine^{2+}

B.1 Experimental details

The experimental details are described in detail elsewhere,[92] thus only a short description is given here. Experiments were performed with a crossed beams device,[188] where a pulsed beam of highly charged Xe^{25+} ions interacts with a beam of neutral glycine molecules produced by evaporation of commercial high purity powder of glycine and glycine-2-2-d_2 in an oven kept at 400K. After the interaction, the cationic products are orthogonally extracted into a linear time-of-flight mass spectrometer. Time-of-flight spectra are recorded in an event-by-event mode allowing to measure the correlation between the charged fragments proceeding from a single ion-molecule collision.[92] Tab.B.1 shows the relative intensities with the corresponding assignation of the correlation islands in the regions of interest in the coincidence map shown in Fig.8.2. Fig.B.1 shows the mass spectrum of the cationic products obtained by fragmentation of the glycine dication in two charged fragments after the interaction of neutral glycine with Xe^{25+} ions at the energy of 387.5 keV.

B.2 Theoretical methods

Quantum chemistry calculations were carried out using the density functional theory (DFT). In particular, geometry optimisations were performed using the B3LYP functional, which combines the Becke's three parameter nonlocal hybrid exchange potential[114] with the nonlocal correlation functional of Lee, Yang and Parr.[215] This functional has been used in combination with the 6-311++G(d,p) basis set. This level of theory has been shown to give accurate results for similar systems.[72] Harmonic vibrational frequencies have been also evaluated at the same level to characterise minima and transition states in the potential energy surface (PES) and to compute the Zero Point Energy (ZPE) correction. For the obtained transition states, intrinsic reaction coordinate (IRC) calculations have been also carried

out to verify the minima they connect. Fig.D.2 in the Chap.8 shows the most important points in the PES that explain the fragmentation paths of doubly-ionised glycine. Figs.B.2, B.3 and B.4 show complementary information obtained on the exploration of the potential energy surface.

Ab initio molecular dynamics (AIMD) calculations were performed using the Atom-centered Density Matrix Propagation method (ADMP),[141;142;149] at the B3LYP/6-31++G(d,p) level of theory. To ensure the adiabaticity during the propagation we chose a time step of $\Delta t = 0.1$ fs and fictitious electron mass of 0.1 amu. We take the geometry of the five most stable conformers of neutral glycine and from them we run simulations increasing the internal vibrational energy (0.001, 0.005, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1, 0.11, 0.12, 0.13, 0.14 and 0.15 Hartrees). The initial velocity of the atoms were randomly selected for each trajectory, isomer and value of the internal energy. A total of 1700 trajectories were performed (20 trajectories for each of the 5 isomers considering 17 values of internal excitation energy). We extract statistics from them by analysing the fragments produced in the simulations at a propagation time of $t_{\max}$ =200 fs. In particular we have observed eight possible exit channels in our simulations:

(a) $Gly^{2+} \rightarrow NH_2CH_2^+ + COOH^+$
(b) $Gly^{2+} \rightarrow NH_3CH^+ + COOH^+$
(c) $Gly^{2+} \rightarrow [NH_2 - CH - C(OH)_2]^{2+}$
(d) $Gly^{2+} \rightarrow [NH_3 - CH - COOH]^{2+}$
(e) $Gly^{2+} \rightarrow NH_2CHCOOH^+ + H^+$
(f) $Gly^{2+} \rightarrow NH_3CHOH^{2+} + CO$
(g) $Gly^{2+} \rightarrow NH_2CHCO^+ + H_2O^+$ / $NH_2CHCO^{2+} + H_2O$
(h) $Gly^{2+} \rightarrow NH_2C^+ + H_2 + COOH^+$

Processes (a) and (b) are Coulomb explosion, (c) and (d) are isomerisation, (e) is H^+ loss, (f) is CO loss, (g) is H_2O or H_2O^+ loss and (h) is a Coulomb explosion combined with a H_2 loss, exit channels. The results of the statistics are shown in Tab.B.2. As a general conclusion, Coulomb explosion (a-b) is observed in 97.94%, isomerisation (c-d) in 1,65% and the rest of the processes (e-h) in 0.41% of the total number of simulations performed.

All the calculations were performed using the Gaussian09 program.[195]

B.3 Results

M1	M2	Relative Intensity	Assignation
18	55	0.05%	H_2O^+/NH_2CHCO^+
18	57	0.14%	$H_2O^+/HNCCO^+$
27	28	1.04%	HNC^+/CO^+
27	29	0.93%	HNC^+/COH^+
27	30	0.36%	HNC^+/COH_2^+
27	45	0.91%	$HNC^+/COOH^+$
28	29	5.20%	$NHCH^+/COH^+$ NH_2CH^+/CO^+
28	30	4.77%	$CO^+/NH_2CH_2^+$ $NHCH^+/COH_2^+$
28	45	7.71%	$NHCH^+/COOH^+$
29	30	2.05%	$COH^+/NH_2CH_2^+$ NH_2CH^+/COH_2^+
29	45	2.37%	$NH_2CH^+/COOH^+$
30	45	3.82%	$NH_2CH_2^+/COOH^+$

Table B.1: Assignation of the correlation islands corresponding to the regions of interest in the coincidence map (see Fig.8.2). Relative intensities are given in percentage of the total intensity of the map.

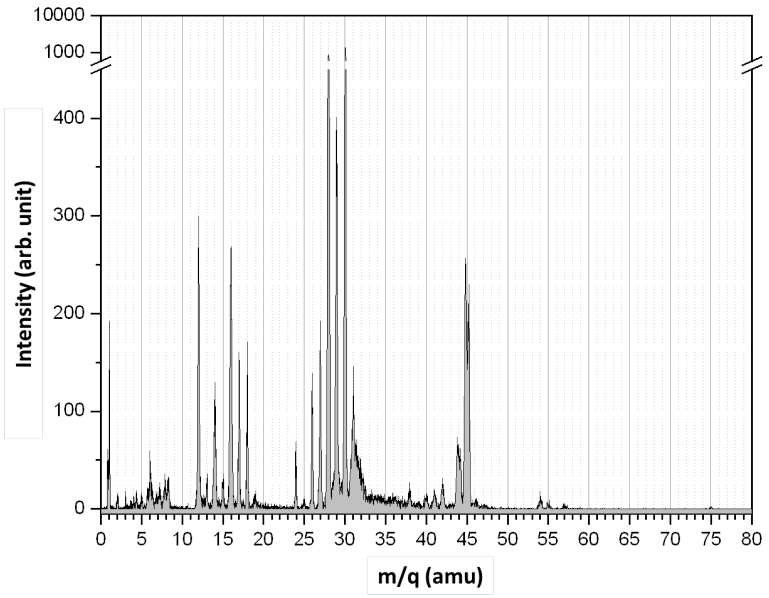


Figure B.1: Mass spectrum of the cationic products obtained by fragmentation of the glycine dication in two charged fragments after the interaction of neutral glycine with Xe^{25+} ions at the energy of 387.5 keV.

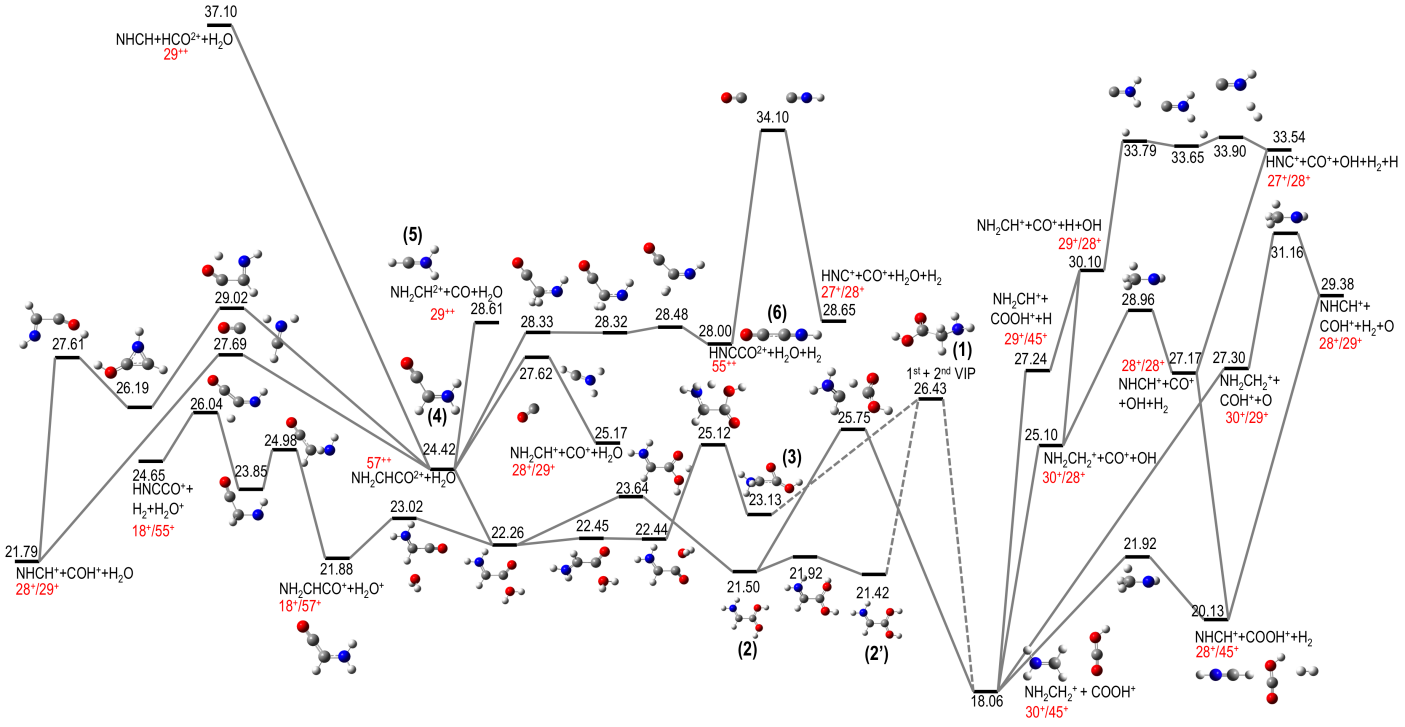


Figure B.2: Most relevant fragmentation pathways for doubly positively charged glycine. Relative energies, ΔE, are given in eV at the DFT-B3LYP/6-311++G(d,p) level of theory including the zero point energy correction with respect to the most stable isomer of the neutral molecule. Key structures as cited in the article are labelled as (1), (2), etc. Coincidence notation is also included, e.g. $30^+/45^+$.
Left part: fragmentation channels leading to doubly charged species after H migration and neutral moieties lost.
Right part: fragmentation channels leading to coincident measurements after Coulomb explosion.
Other channels are shown in Fig.B.3, isomerisation of doubly charged glycine, and in Fig.B.4, Coulomb explosion after H migration.

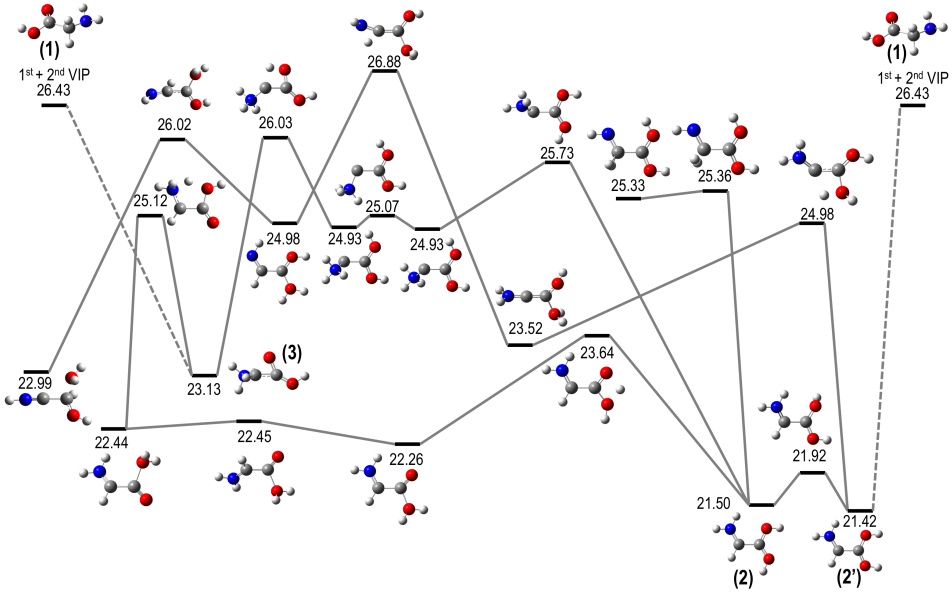


Figure B.3: Isomerisation pathways for doubly positively charged glycine. Relative energies, ΔE, are given in eV at the DFT-B3LYP/6-311++G(d,p) level of theory including the zero point energy correction with respect to the most stable isomer of the neutral molecule. Key structures as cited in the article are labelled as (1), (2), etc. Notice that the atomic arrangement of the initial glycine molecule, NH_2CH_2COOH does not exist as stable isomer in dicationic glycine: geminal diol $-C(OH)_2$ (obtained from (1) in the right part of the figure) in and/or $-NH_3$ terminations (obtained from (1) in the left part of the figure) stabilise the charge−charge repulsion.

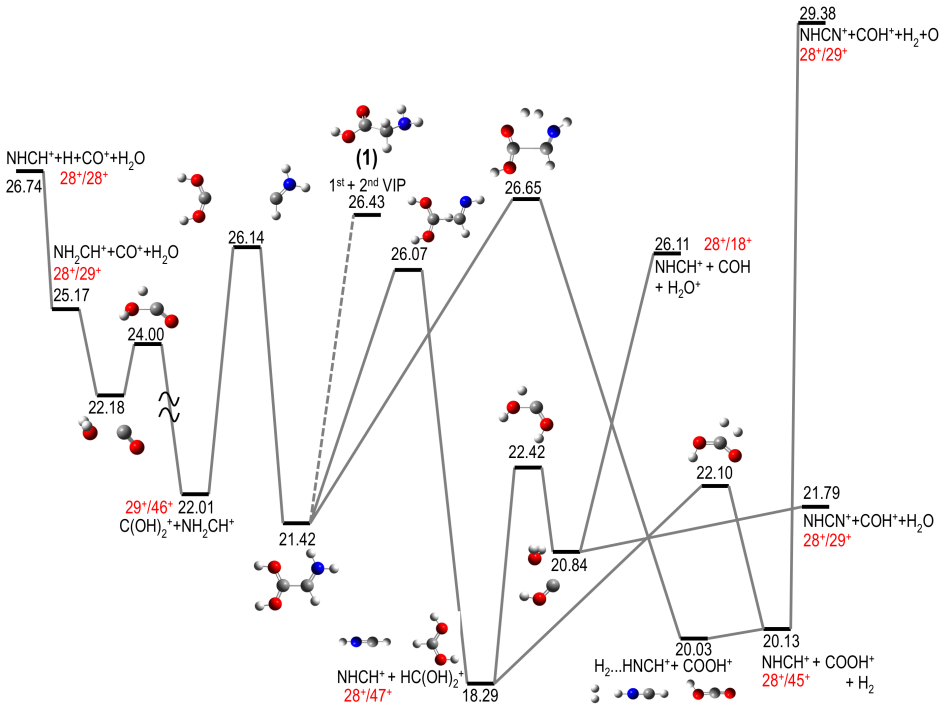


Figure B.4: Fragmentation pathways for doubly positively charged glycine. Channels leading to coincident measurements after Coulomb explosion from the geminal diol isomer are shown. Relative energies, ΔE, are given in eV at the DFT-B3LYP/6-311++G(d,p) level of theory including the zero point energy correction with respect to the most stable isomer of the neutral molecule. Key structures as cited in the article are labelled as (1), (2), etc. Coincidence notation is also included, e.g. $30^+/45^+$. In this figure, the fragmentation pathways after ultrafast H migration leading to to coincident measurements are given.

Energy (eV)	% Percentage of each process							
	Coulomb Explosion		Isomerisation		Others			
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
0.027	100%	0%	0%	0%	0%	0%	0%	0%
0.136	100%	0%	0%	0%	0%	0%	0%	0%
0.272	100%	0%	0%	0%	0%	0%	0%	0%
0.544	100%	0%	0%	0%	0%	0%	0%	0%
0.816	100%	0%	0%	0%	0%	0%	0%	0%
1.089	100%	0%	0%	0%	0%	0%	0%	0%
1.361	100%	0%	0%	0%	0%	0%	0%	0%
1.633	99%	0%	0%	1%	0%	0%	0%	0%
1.905	98%	0%	0%	2%	0%	0%	0%	0%
2.177	98%	0%	0%	2%	0%	0%	0%	0%
2.449	98%	1%	1%	0%	0%	0%	0%	0%
2.722	98%	0%	1%	1%	0%	0%	0%	0%
2.994	94%	2%	1%	1%	1%	1%	0%	0%
3.266	92%	4%	0%	4%	0%	0%	0%	0%
3.538	91%	5%	0%	4%	0%	0%	0%	0%
3.810	91%	3%	0%	5%	0%	0%	1%	0%
4.082	86%	5%	1%	4%	1%	1%	1%	1%
Total – average	96.76%	1.18%	0.24%	1.41%	0.12%	0.12%	0.12%	0.06%
	97.94%		1.65%		0.42%			

Table B.2: Statistics for each value of the excitation energy, summed over all the considered isomers, percentage of the observed exit channels in the molecular dynamics simulations: (a) $NH_2CH_2^+ + COOH^+$, (b) $NH_3CH^+ + COOH^+$, (c) $[NH_2 - CH - C(OH)_2]^{2+}$, (d) $[NH_3 - CH - COOH]^{2+}$, (e) $NH_2CHCOOH + +H^+$, (f) $NH_3CHOH^{2+} + CO$, (g) $NH_2CHCO^+ + H_2O^+$ / $NH_2CHCO^{2+} + H_2O$, (h) $NH_2C^+ + H_2 + COOH^+$.

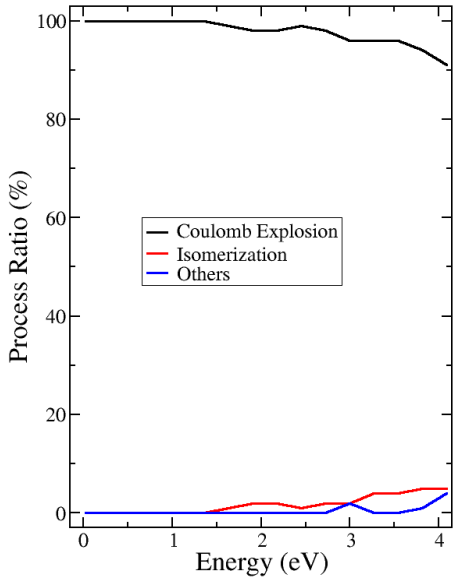


Figure B.5: Percentage of processes as a function of the internal excitation energy: Coulomb explosion processes (a) and (b), Isomerisation processes (c) and (d), and Other processes (e-h).

Appendix C

Appendices
Additional information for β-alanine^{2+}

C.1 Detailed statistics

Detailed statistics of our *ab initio* molecular dynamics simulations are shown in Fig. C.1, C.2 and C.3. Total contribution of direct bond breaking in a Coulomb repulsion process is presented in Fig. C.1. We can see that $C_{carboxyl} - C_\alpha$ dominates over $C_\alpha - C_\beta$ bond breaking in the whole range of the excitation energies. Fig. C.2 shows the statistical results of the isomerisation of β-alanine leading to diol^{2+} (5), (4-membered ring)$^{2+}$ (8), (5-membered ring)$^{2+}$ (6) and linear $NH_3CH_2CHCOOH^{2+}$ (7) as a function of the excitation energy. These channels are much less populated, about 5% (200 trajectories) of the total simulations. Isomerisation to diol^{2+}, which is the most probable isomerisation mechanism, is mainly observed for lower values of excitation energy. For larger excitation energies, we observe competition

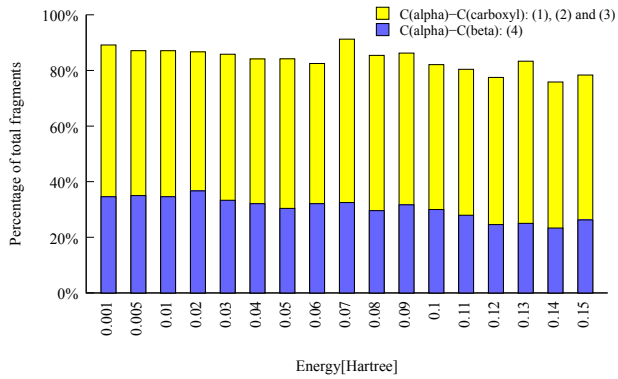


Figure C.1: Statistics of the AIMD simulations showing Coulomb explosion: competition between $C_{carboxyl} - C_\alpha$: processes (1), (2), (3) and $C_\alpha - C_\beta$ bond breaking: process (4).

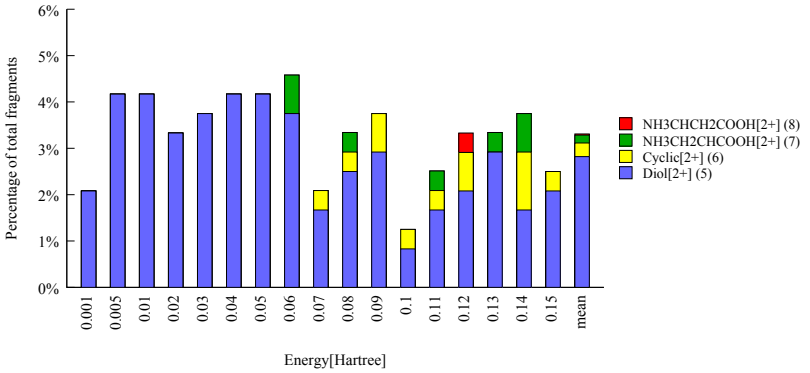


Figure C.2: AIMD simulations showing isomerisation competition.

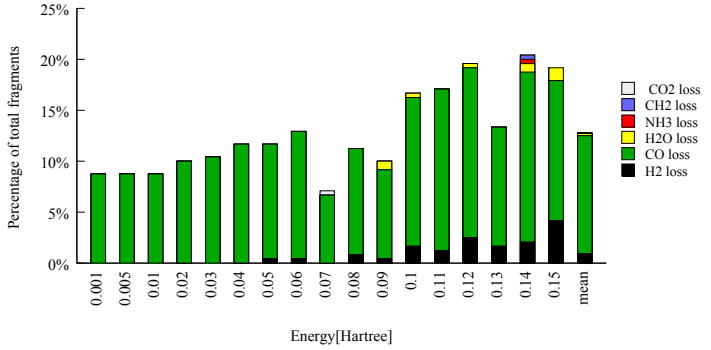


Figure C.3: AIMD simulations showing combined processes competition.

with the other processes. Fig. C.3 shows combined processes. The probability of occurrence of the combined processes increases with the excitation energy. It is not surprising, as the additional excitation energy given to the vibrational degrees of freedom can induce further molecular rearrangements. Therefore, additional rearrangements appear together with the expected Coulomb repulsion, thereby decreasing the presence of repulsion processes where the excitation energy increases.

C.2 Conformer studies

Conformers nomenclature follows the work by Sanz et al.[55] The fragmentation pattern depends on geometry of each isomer (see Fig. C.5, Fig. C.6 and Fig. C.7) and is classified into four groups:

- $NH_2CHCH_3^+ + COOH^+$ (1): isomers a1,a2, a6, a11

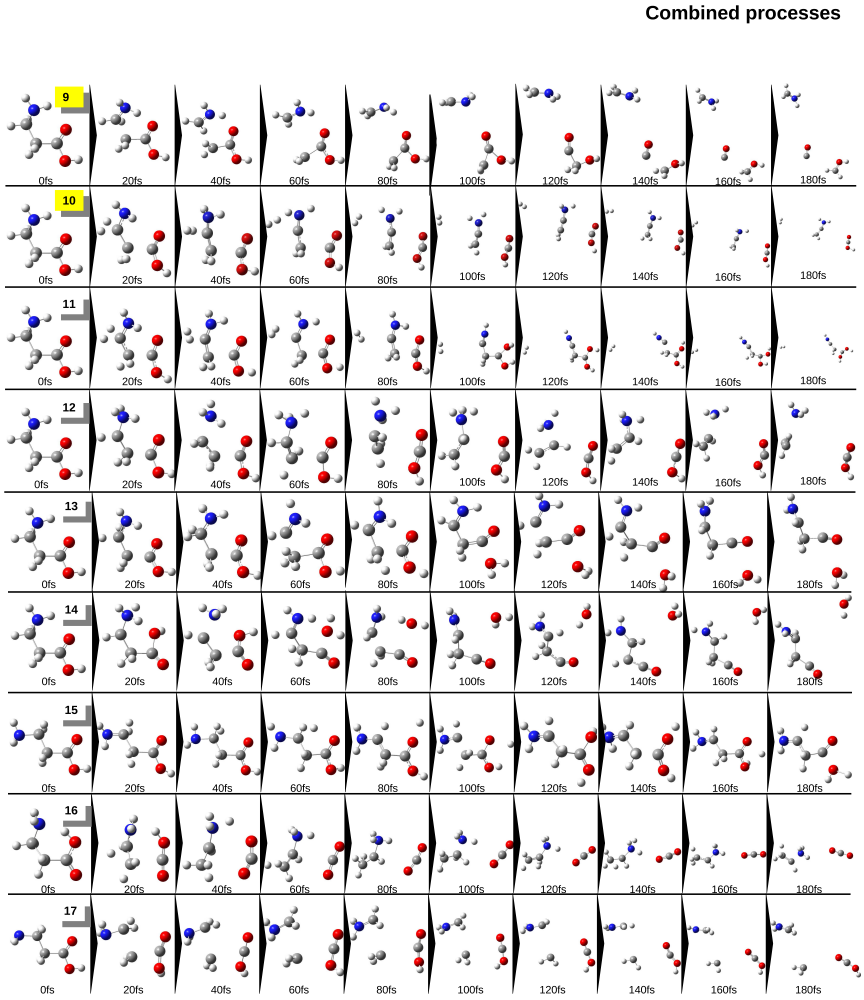


Figure C.4: Snapshots of molecular dynamics simulations of a β-alanine dication giving though a combined processes: a) Coulomb explosion, isomerisation and fragmentation: (9); b) Coulomb explosion and fragmentation: (10), (17); c) fragmentation and isomerisation: (11); d) isomerisation, Coulomb explosion and fragmentation: (12); e) isomerisation and dehydration: (13), (14), (15); f) isomerisation and Coulomb explosion: (16).

- $NH_2CHCH_3^+ + COOH^+$ (1) in competition with $NH_2CH_2^+ + CH_2OH^+ + CO$ (9): isomers a3 and a12

- $NH_2CH_2^+ + CH_2COOH^+$ (4): isomers a7, a8, a9, a10

- $NH_2CH_2^+ + CH_2OH^+ + CO$ (9) in competition with $NH_2NH_2NH_2^+ + COOH^+$ (3): isomer a5

The isomer a4 can fragment in many different ways and it is not classified in any of the group above. Its geometry is very different from any other isomer and slightly similar to the geometry of a5. Indeed, conformer a5 also present a different fragmentation behaviour than the others.

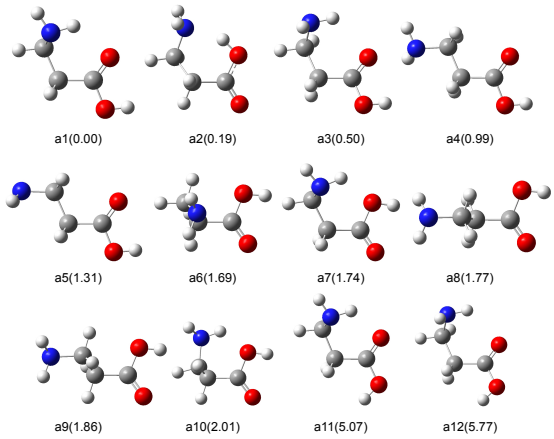


Figure C.5: Geometries of the most stable 12 neutral conformers of beta-alanine. The numbers in brackets indicate the relative energy between them in kcal mol^{-1}.

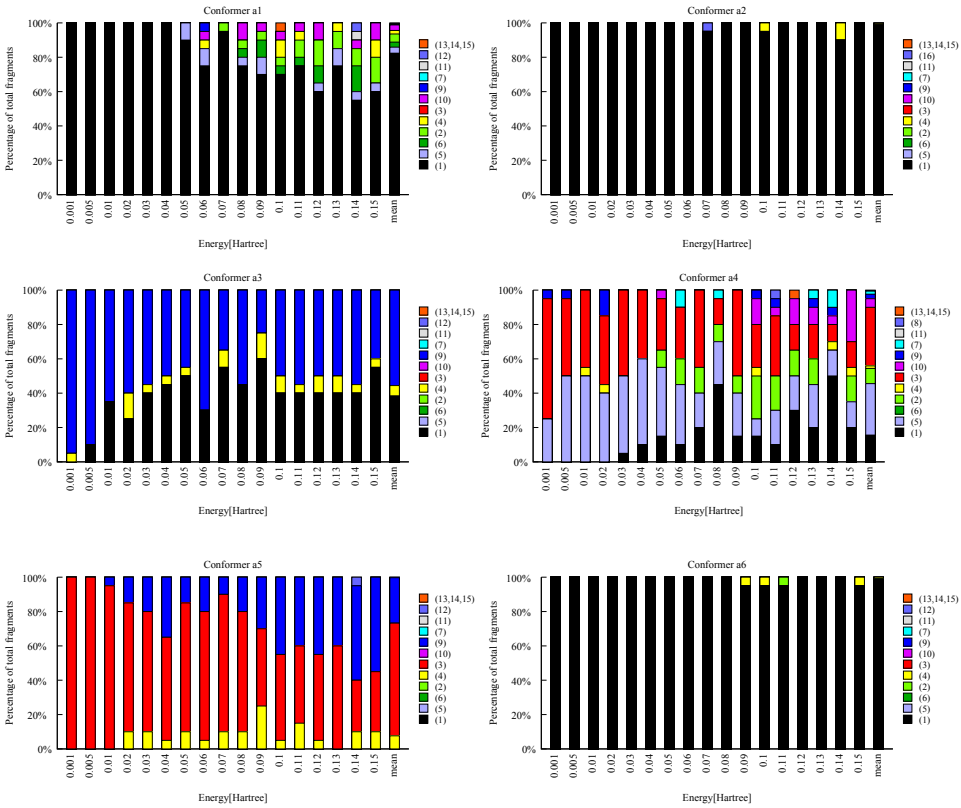


Figure C.6: *Ab initio* molecular dynamics statistics for the six most stable conformers (a1-a6) of beta-alanine with relative energy between them 1.7 kcal mol $^{-1}$. Numbers in brackets correspond to the mechanisms defined in the article and Fig. 4.

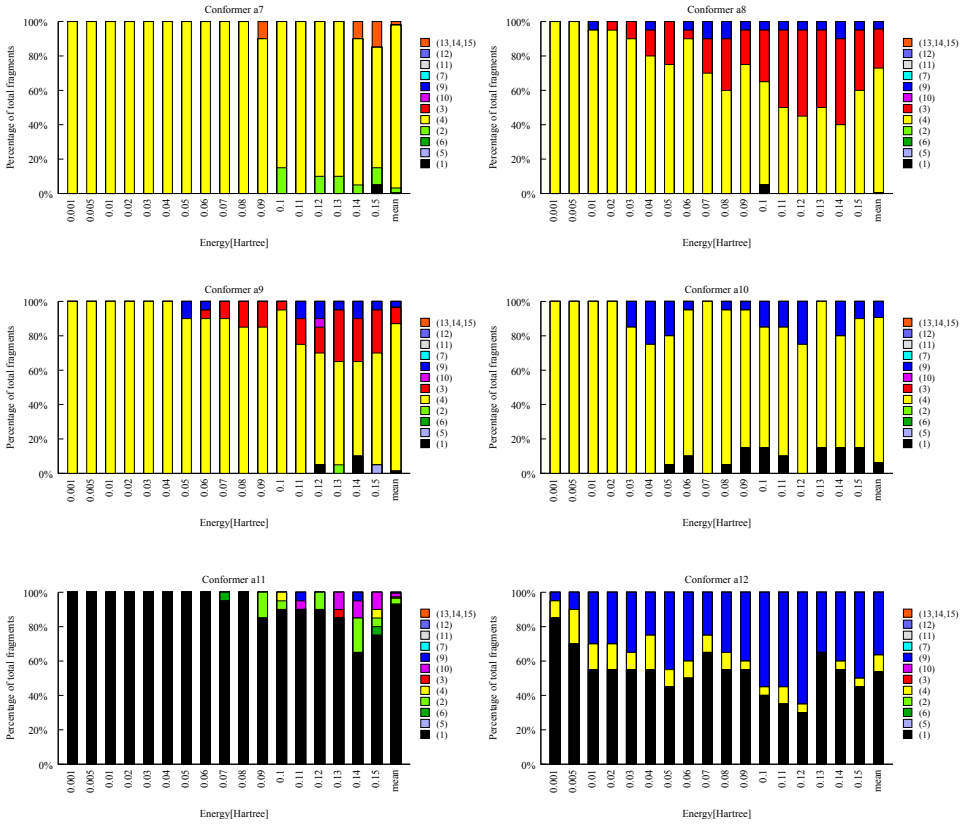


Figure C.7: *Ab initio* molecular dynamics statistics for six conformers (a7-a12) of beta-alanine with relative energy between them 5.8 kcal mol $^{-1}$. Numbers in brackets correspond to the mechanisms defined in the article and Fig. 4.

C.3 Slope determination

The correlation between two charged fragments detected in coincidence gives information about the fragmentation dynamics of the doubly charged molecules. The coincidence map displays the time-of-flight of the slower cationic fragment (TOF2) as a function of the time-of-flight of the faster cationic fragment (TOF1). The correlation for a given pair of charged fragments corresponds to an island on the coincidence map. The width in TOF is proportional to the momentum. The slope of islands in coincidence map can be related to the momentum of the respective cationic fragments. In a deeper analysis, the width in the TOF spectrum and the slope of one island in the coincidence map can be related to the momentum (p_i) of

the respective cationic fragments:

$$S = -\frac{\Delta TOF_2}{\Delta TOF_1} = -\frac{p_2}{p_1} \tag{C.1}$$

, where ΔTOF_1 and ΔTOF_2 are the width in the TOF of the lighter and heavier fragments, respectively.
If the two cations stem from a two-body break-up or if there is a loss of a neutral product with no kinetic energy released before fragmentation of the dication into two charged species, conservation of momentum implies a slope of -1 in the respective island. In the case of a secondary decay mechanism, the island has a steeper or gentle slope. Thus, comparing the experimental slope, S_{exp}, from the shape of the island and the theoretical slope, S_{theor}, calculated from the proposed mechanisms, we can infer where the charged fragments originally come from. Experimental slopes were plotted for the most intense islands on the coincidence map. Therefore, theoretically calculated channels and experimentally measured slopes shed light on the chemical reactions and explain the behaviour in the fragmentation dynamics. Comparison between the calculated slopes obtained from the computed mechanisms with the experimental measurements with the associated channels are presented in C.1.

C.4 Channels

Fig.C.8 summarises the different studied channels, from the main pathways observed in the molecular dynamics simulations: $C_{carboxyl} - C_\alpha$ bond breaking, $C_\alpha - C_\beta$ bond breaking, H migration and OH migration; and the subsequent exploration of the potential energy surface.

Table C.1: Assignation of the correlation islands corresponding to the regions of interest in the coincidence map for β-alanine dication. I_{exp} is the relative intensity as measured in the experiment. E is the highest energy barrier needed to achieve this exit channel. The number in brackets indicate the mechanism as given in the molecular dynamics simulations: (1) $NH_2CHCH_3^+ + COOH^+$, (2) $NH_3CHCH_2^+ + COOH^+$, (4) $NH_2CH_2^+ + CH_2COOH^+$, (5) $NH_2CHCH_2CO^{2+} + H_2O$, (9) $NH_2CH_2^+ + CH_2OH^+ + CO$ in the assignation column comes from $C_{carboxyl} - C_\alpha$ (1 and 2), $C_\alpha - C_\beta$ (4), H migration (5) and OH migration processes (9), respectively.

M1	M2	I_{exp}	Assignation			E	S_{theor}	S_{exp}
30	42	100	(4)	$NH_2CH_2^+ + CH_2COOH^+$	$\rightarrow$ $NH_2CH_2^+ + CH_2CO^+ + OH$	21.15	-0.71	-1.0
				$NH_2CH_2CH_2CO^{2+} + OH$	$\rightarrow$ $NH_2CH_2^+ + CH_2CO^+ + OH$	25.70	-1.00	
30	31	80	(9)	$NH_2CH_2^+ + CH_2COOH^+$	$\rightarrow$ $NH_2CH_2^+ + CH_2OH^+ + CO$	19.55	-0.60	-
42	45	58	(1)	$NH_2CHCH_3^+ + COOH^+$	$\rightarrow$ $NH_2CCH_2^+ + H_2 + COOH^+$	20.62	-1.05	-1.1
			(1)	$NH_2CHCH_3^+ + COOH^+$	$\rightarrow$ $NHCCH_3^+ + H_2 + COOH^+$	21.33	-1.05	
28	42	47	(4)	$NH_2CH_2^+ + CH_2COOH^+$	$\rightarrow$ $NHCH^+ + H_2 + CH_2CO^+ + OH$	25.01	-0.76	-1.0
			(1)	$NH_2CHCH_3^+ + COOH^+$	$\rightarrow$ $NHCCH_3^+ + H_2 + CO^+ + OH$	25.95	-1.53	
			(5)	$NH_2CHCH_2CO^{2+} + H_2O$	$\rightarrow$ $NH_2CCH_2^+ + H + CO^+ + H_2O$	26.04	-0.98	
30	45	45	(4)	$NH_2CH_2^+ + CH_2COOH^+$	$\rightarrow$ $NH_2CH_2^+ + CH_2 + COOH^+$	22.60	-0.76	-1.4
			(1)	$NH_2CHCH_3^+ + COOH^+$	$\rightarrow$ $NH_2CH_2^+ + CH_2 + COOH^+$	22.60	-1.47	
28	45	39	(4)	$NH_2CH_2^+ + CH_2COOH^+$	$\rightarrow$ $NHCH^+ + H_2 + CH_2 + COOH^+$	24.68	-0.82	-1.4
			(1)	$NH_2CHCH_3^+ + COOH^+$	$\rightarrow$ $NHCH^+ + CH_2 + H_2 + COOH^+$	26.46	-1.57	
28	29	35	(4)	$NH_2CH_2^+ + CH_2COOH^+$	$\rightarrow$ $NHCH^+ + H_2 + CH + COH^+ + OH$	30.59	-0.53	-
28	30	31	(4)	$NH_2CH_2^+ + CH_2COOH^+$	$\rightarrow$ $NH_2CH_2^+ + CH_2OH + CO^+$	24.72	-2.11	-
43	45	30	(1)	$NH_2CHCH_3^+ + COOH^+$	$\rightarrow$ $NH_2CHCH_2^+ + H + COOH^+$	21.17	-1.02	-1.1
			(1)	$NH_2CHCH_3^+ + COOH^+$	$\rightarrow$ $NH_2CCH_3^+ + H + COOH^+$	22.01	-1.02	
			(1)	$NH_2CHCH_3^+ + COOH^+$	$\rightarrow$ $NHCHCH_3^+ + H + COOH^+$	23.91	-1.02	
29	42	29	(4)	$NH_2CH_2^+ + CH_2COOH^+$	$\rightarrow$ $NH_2CH^+ + H + CH_2CO^+ + OH$	26.14	-0.74	-1.0
			(4)	$NH_2CH_2^+ + CH_2COOH^+$	$\rightarrow$ $NHCCH_3^+ + H_2 + COH^+ + O$	28.14	-1.48	
			(4)	$NH_2CH_2^+ + CH_2COOH^+$	$\rightarrow$ $NH_2CCH_2^+ + H_2 + COH^+ + O$	28.67	-1.48	
			(5)	$NH_2CHCHCOH^{2+} + H_2O$	$\rightarrow$ $NH_2CH^+ + CHCOH^+ + H_2O$	27.05	-1.00	
			(5)	$NH_2CHCH_2CO^{2+} + H_2O$	$\rightarrow$ $NH_2CH^+ + CH_2CO^+ + H_2O$	25.45	-1.00	
			(5)	$NH_2CHCH_2CO^{2+} + H_2O$	$\rightarrow$ $NH_2CHCH^+ + COH^+ + H_2O$	21.09	-1.00	
			(5)	$NH_2CHCHCOH^{2+} + H_2O$	$\rightarrow$ $NH_2CHCH^+ + COH^+ + H_2O$	21.09	-1.00	
29	30	28	(4)	$NH_2CH_2^+ + CH_2COOH^+$	$\rightarrow$ $NH_2CH_2^+ + CH + COH^+ + OH$	28.52	-2.03	-
41	45	22	(1)	$NH_2CHCH_3^+ + COOH^+$	$\rightarrow$ $NHCCH_2^+ + H_2 + H + COOH^+$	23.06	-1.07	-1.1
			(1)	$NH_2CHCH_3^+ + COOH^+$	$\rightarrow$ $NH_2CCH^+ + H_2 + H + COH^+$	23.76	-1.07	
			(1)	$NH_2CHCH_3^+ + COOH^+$	$\rightarrow$ $NCCH_3 + H_2 + H + COOH^+$	25.84	-1.07	
14	28	21	(4)	$NH_2CH_2^+ + CH_2COOH^+$	$\rightarrow$ $NHCH^+ + H_2 + CH_2^+ + COOH$	26.32	-3.93	-
18	45	19	(2)	$NH_3CHCH_2^+ + COOH^+$	$\rightarrow$ $NH_4^+ + CHCH + COOH^+$	20.74	-2.44	-
27	45	19	(2)	$NH_3CHCH_2^+ + COOH^+$	$\rightarrow$ $NH_3 + CHCH_2^+ + COOH^+$	21.50	-1.63	-1.5
15	45	17	(1)	$NH_2CHCH_3^+ + COOH^+$	$\rightarrow$ $NH_2CH + CH_3^+ + COOH^+$	24.04	-2.93	-
28	31	14	(9)	$NH_2CH_2^+ + CH_2OH^+ + CO$	$\rightarrow$ $NHCH^+ + H_2 + CH_2OH^+ + CO$	22.01	-0.56	-
			(2)	$NH_3CHCH_2^+ + COOH^+$	$\rightarrow$ $NH_3CH_2^+ + CH + CO^+ + OH$	32.19	-1.16	

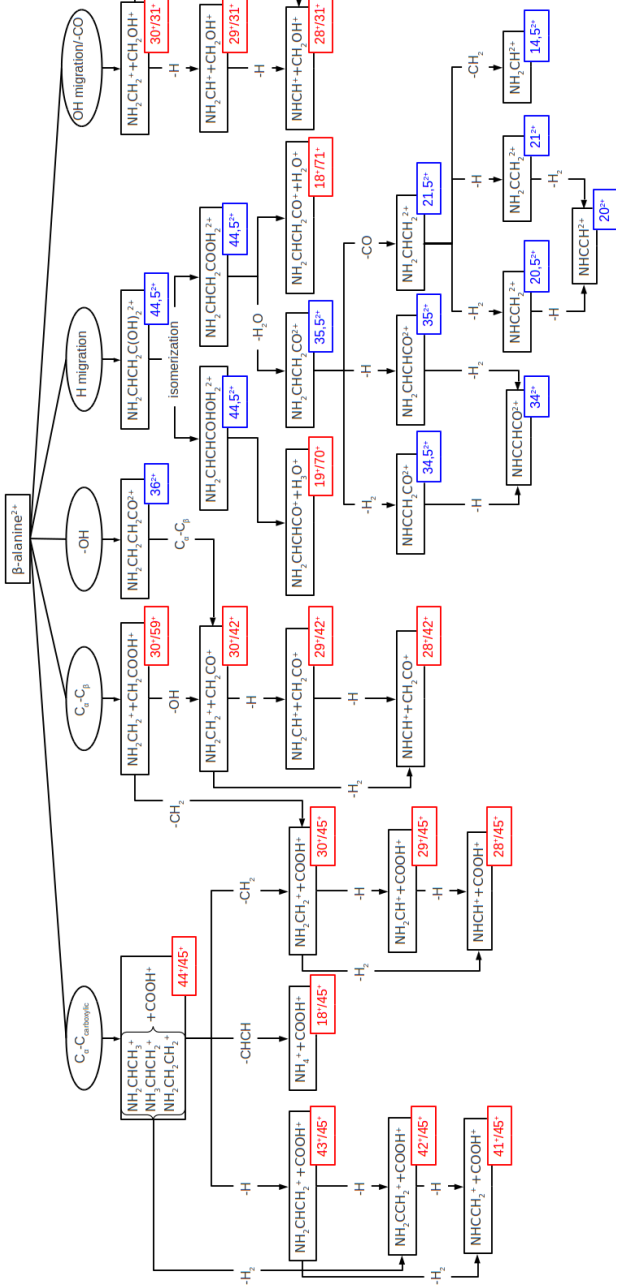


Figure C.8: Schematic picture describing the five most probable fragmentation channels of doubly charged β-alanine observed in the ADMP simulations. Subsequent dissociation pathways are also presented with coincidence notation, *e.g.* $44^+/45^+$.

Appendices

ADDITIONAL INFORMATION FOR $GABA^{2+,3+}$

D.1 Double electron capture.

Detailed statistics of *ab initio* molecular simulations (AIMD) related with Coulomb explosion (CE) of $GABA^{2+}$ molecules are shown in Fig. D.1a). The fragmentation yield of CE processes: (1), (2), (3), and (4) is presented as a function of excitation energy. In Fig. D.1b) we also present the optimised fragments taken from MD simulations with their relative energy (in parenthesis); mechanisms are associated according to the classification of the bond cleavage. We can also see that the probability of the CE processes (3) and (4) is much higher than for other CE types. $C_{carboxyl} - C_{\alpha}$ bond cleavage (1) starts to increase with the excitation en-

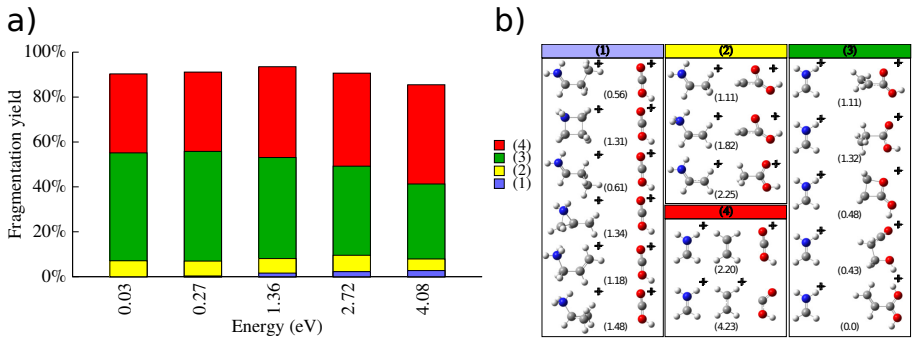


Figure D.1: *Ab initio* molecular dynamics statistics for all 30 dications of GABA. Competition between different Coulomb bond cleavages. The number in parentheses represents different bonds breaking: (1) $C_{carboxyl} - C_{\alpha}$, (2) $C_{\alpha} - C_{\beta}$, (3) $C_{\beta} - C_{\gamma}$ and (4) $C_{carboxyl} - C_{\alpha} + C_{\beta} - C_{\gamma}$ bond cleavages; a) fragmentation yield as a function of the excitation energy (in eV); b) optimised MD exit channels with DFT-B3LYP/6-311G(d,p) level of theory, where the numbers in brackets indicate the relative energy in respect with the most stable fragmentation channel in eV.

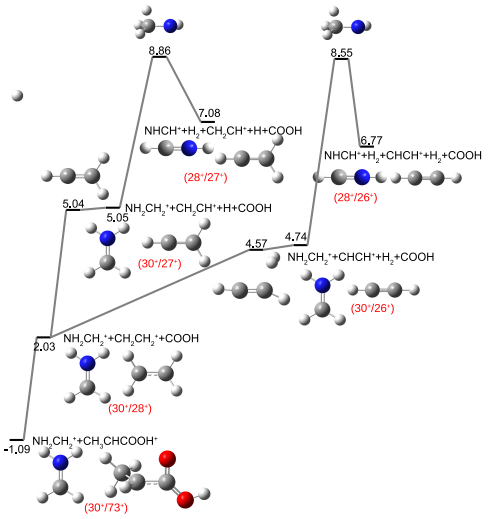


Figure D.2: Fragmentation pathways for doubly positively charged GABA. Only the stationary points of the potential energy surface corresponding to $C_\beta - C_\gamma$ bond breaking and further hydrogen emission leading to charged fragments: m/z=28 amu with other heavy fragments (m/z=26 and 27 amu). Relative energies, ΔE, are given in eV at the DFT-B3LYP/6-311++G(d,p) level of theory including the zero point energy correction with respect to the exit channel: $NH_2CH_2^+ + CH_2CH_2 + COOH^+$, which is exothermic with respect to GABA^{2+} by 5.28 eV (see Tab. 1 in main article).

ergy but even for 4 eV of excitation energy it is still significantly less probable than other CE mechanisms. Interestingly, in Fig. D.1b) we observe 3 and 4-member ring structures after $C_{carboxyl} - C_\alpha$ and $C_\beta - C_\gamma$ bond cleavages (channels (1) and (3), respectively). Considering three body fragmentation of GABA^{2+} (process (4) in Fig. D.1) charge could be hypothetically distributed in six different ways: a) $NH_2CH_2^+ + CH_2CH_2 + COOH^+$, b) $NH_2CH_2^+ + CH_2CH_2^+ + COOH$, c) $NH_2CH_2 + CH_2CH_2^+ + COOH^+$, d) $NH_2CH_2^{2+} + CH_2CH_2 + COOH$, e) $NH_2CH_2 + CH_2CH_2^{2+} + COOH$ or f) $NH_2CH_2 + CH_2CH_2 + COOH^{2+}$. The MD simulations reveal that among them, only a) and b) are noticed with 92.7% and 7.3% probabilities, respectively. Total contribution of the channels 4a and 4b to the MD statistics for double electron capture is 32.4% and 2.5%, respectively. Moreover, the former exit channel (4a) is 2 eV lower in energy. We also see that $C_\beta - C_\gamma$ (3) dominates over another 2 body fragmentation in the whole range of excitation energies. One would expect that in the longer simulation timescale the (3) process will contribute to the (4) due to the further $C_{carboxylic} - C_\alpha$ bond breaking of the CH_2COOH^+ ending up either on $CH_2CH_2 + COOH^+$ (2.20 eV) or $CH_2CH_2^+ + COOH$ (4.23 eV).

Two intense coincidence islands $26^+/28^+$ and $27^+/28^+$ related with CE processes are analysed in details on the potential energy surface (PES) (see Fig. D.2) in order to gain a better understanding of the experimental measurements. We

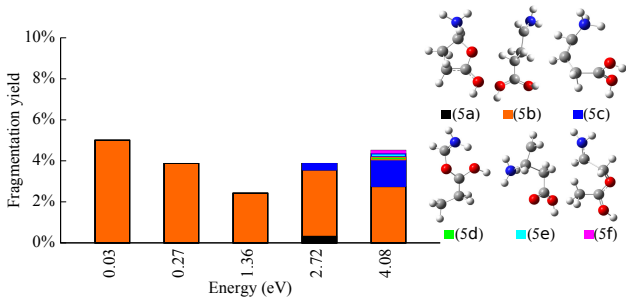


Figure D.3: Statistics of the AIMD simulations showing isomerisation competition for $GABA^{2+}$. The number in parentheses represents different geometry reorganisation of dication of GABA taken from the MD simulation and than optimised at the DFT-B3LYP/6-311++G(d,p) level of theory.

observe, not only similar relative energies (6.77 and 7.08 eV) but also similar energy barriers. It explain similar experimental relative intensities measured in the experiment (3.2% and 2.5%). Even if the relative theoretical and experimental comparison between both channels are in an agreement we see that both channels are much higher in the energy than the ones presented in the Tab. 1 in main article and involve high energy barriers on PES. However, both pathways can be derived from MD simulations (3) and (4) after neutral H and H_2 emissions as a kinetically favourable processes. This emphasises the role of the molecular dynamics simulations to properly describe fragmentation.

The degeneracy of another exit channels: $NHCH^+ + H_2 + CH_2CH_2 + COOH^+$ and $NH_2CH_2^+ + CH_2CH_2^+ + COOH$ explains the similar experimental (relative intensity) and theoretical (relative energy) values (see Tab. 1 in main article).

In the Fig. D.3 we see results of MD simulation for $GABA^{2+}$, without bond breaking; H-migration leads to stable isomers (denoted as (5a-f)). The highest yield is observed for diol $^{2+}$ isomers.

Fig. D.4 shows combined processes for $GABA^{2+}$ which are related with the emission of neutrals before or after CE and isomerisation. The probability of occurrence of the combined processes increases with the excitation energy (see Fig. D.4). It is not surprising, as the additional excitation energy given to the vibrational degrees of freedom can induce many secondary reactions.

Possible structures corresponding to doubly charged fragments (m/z = 26.5, 27.5, 28.5 amu) of reactive nitrogen species (RNS) observed in experiment and extensively discussed in main article are presented Fig. D.5. The structure of cyclic bigger doubly charged moieties are shown in Fig. D.6. Reactive oxygen species (ROS) with the same masses as RNS described above, proposed to be formed in the fragmentation of $GABA^{3+}$ are presented in Fig. D.10. Some of those structures are obtained after the optimisation of neutral molecules taken from the NIST database,[391] the rest is obtained after optimisation of the fragments observed

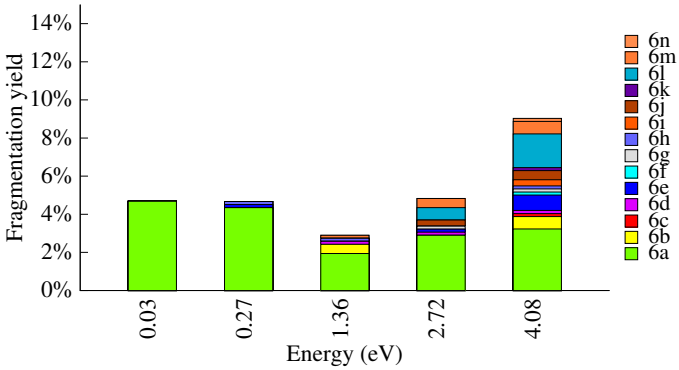


Figure D.4: AIMD simulations showing combined processes competition for $GABA^{2+}$.

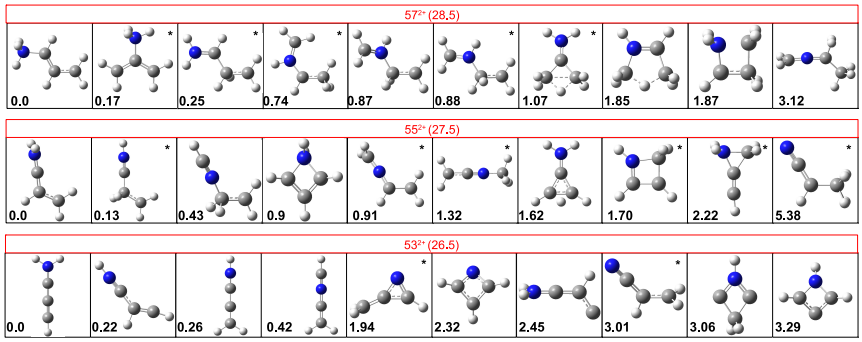


Figure D.5: Possible structures of doubly charged cations at m/z = 26.5, 27.5 and 28.5 amu optimised at the DFT-B3LYP/6-311++G(d,p) level of theory. Relative energies in parentheses between the conformers are in eV. Fragments with star (*) were optimised of the corresponding neutral structure taken from National Institute of Standards and Technology (NIST).

in the MD. According to previous study of the absorption and fluorescence spectra of microporous polymers,[392] even if the ring structures are not in their lowest energy conformation we can expect them to play an important role.

D.2 Triple electron capture

Total contribution of direct bond breaking for $GABA^{3+}$ related with CE is presented in Fig. D.7. The CE processes are very similar in the full excitation energy range to $GABA^{2+}$ fragmentation discussed previously, (see Fig.D.7) considering that the $C_\beta - C_\gamma$ become the dominant pathway for three electron capture phe-

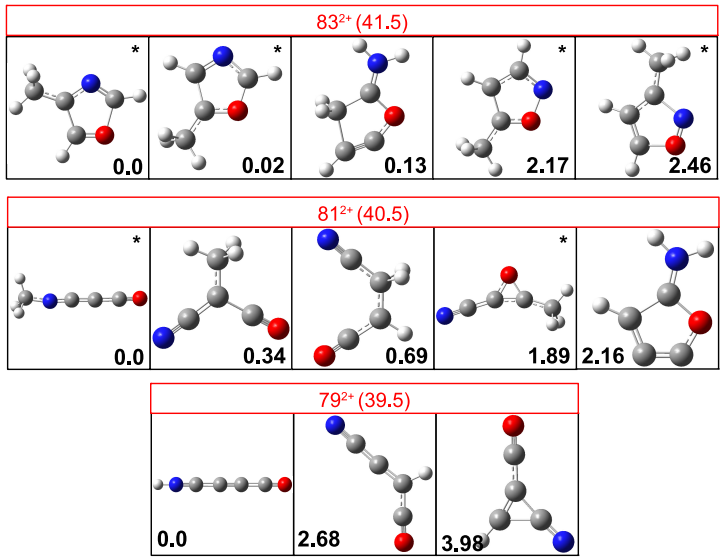


Figure D.6: Possible structures of doubly charged cations at m/z = 39.5, 40.5 and 41.5 amu optimised at the DFT-B3LYP/6-311++G(d,p) level of theory. Relative energies in parentheses between the conformers are in eV. Fragments with star (*) were optimised of the corresponding neutral structure taken from NIST.

nomena and in the slightly longer simulation timescale will break in three fragments similar to $GABA^{2+}$.

MD simulation for $GABA^{3+}$ without bond breaking are presented in Fig. D.8. H-migration leads to isomers denoted as (5a-e). The highest yield is observed for $[NH_3 - R]^{3+}$ forms, where R is the rest of the molecule without one hydrogen.

Fig. D.9 shows combined processes for $GABA^{3+}$. The fragmentation yield of the combined processes increases with the excitation energy similarly to $GABA^{2+}$. The most important processes classified as combined for $GABA^{3+}$ fragmentation correspond to emission of H_2 molecules (6e and 6f) with subsequent $C_{carboxylic} - C_\alpha$ bond cleavage i.e. $COOH^+$ emission (6b) and (6y) $NH_2CCH_3^+ + CH_3^+ + COOH^+$.

The possible optimised reactive oxygen species (ROS) are presented in Fig. D.10. These doubly charged moieties are produced only in the fragmentation of $GABA^{3+}$ due to different double charge localisation in comparison with $GABA^{2+}$.

Fig. D.11 presents the analysis of the charges on the emitted hydrogen molecules related with combined processes. The Mulliken charges for H_2, as it is shown in the panel a) oscillates as a number of counts (bin of 0.01). Panel b) present the counts for H_2 charges considering the threshold smaller than the value of the charge (0 to 1 with 0.01 bin). The plots shows that the charge state of hydrogen molecules vary depending on the criteria. There is no clear neutral (q=0) or charged (q=1) H_2 molecule emission. We consider emission of charged hydrogen molecules if q>0.9,

Table D.1: All 37 mechanisms for the fragmentation of $GABA^{2+}$ molecules.

$C_{carboxylic} - C_\alpha$ (1)	$C_\alpha - C_\beta$ (2)	$C_\beta - C_\gamma$ (3)	$C_{carboxylic} - C_\alpha + C_\beta - C_\gamma$ (4)	Isomerisation (5)	Neutral emission (6)
$NH_2CHCH_2CH_3^+ + COOH^+$ (1a)	$NH_2CHCH_3^+ + CH_2COOH^+$ (2a)	$NH_2CH_2^+ + CH_3CHCOOH^+$ (3a)	$NH_2CH_2^+ + CH_2CH_2 + COOH^+$ (4a)	$NH_3CHCH_2CH_2COOH^{2+}$ (5a)	$NH_2CH_2^+ + CH_2CH_3^+ + CO_2$ (6a)
$NH_2CH_2CH_2CH_2^+ + COOH^+$ (1b)	$NH_3CHCH_2^+ + CH_2COOH^+$ (2b)	$NH_2CH_2^+ + CH_2CH_2COOH^+$ (3b)	$NH_2CH_2^+ + CH_2CH_2^+ + COOH$ (4b)	$NH_2CHCH_2CH_2C(OH)_2^{2+}$ (5b)	$NH_3CH_2CHCH_3^{2+} + CO_2$ (6b)
$NH_2CHCH_2CH_3^+ + COOH^+$ (1c)	$NH_2CHCH_2^+ + CH_3COOH^+$ (2c)	$NH_2CH_2^+ + CH_2(CH_2O)COH^+$ (3c)		$NH_3CHCHCH_2C(OH)_2^{2+}$ (5c)	$NH_3CH_2CH_2CH_2^{2+} + CO_2$ (6c)
$NH_2(CH_2CH)CH_3^+ + COOH^+$ (1d)		$NH_2CH_2^+ + CH_2(CH_2CO)OH^+$ (3d)		$NH_2CHOCOHCH_2CH_3^{2+}$ (5d)	$NH_2CH_2^+ + CH_3CHOH^+ + CO$ (6d)
$NH_3CH_2CHCH_2^+ + COOH^+$ (1e)		$NH_2CH_2^+ + CH_2CHC(OH)_2^+$ (3e)		$CH_2CH(NH_3)CH_2COOH^{2+}$ (5e)	$NH_2CHCH_3^+ + CH_2OH^+ + CO$ (6e)
$NH_3CH(CH_2)_2^+ + COOH^+$ (1f)		$NH_2CH_2^+ + HOCH_2CH_2CO^+$ (3f)		$NH_2CHCH_2OCCH_2OH^{2+}$ (5f)	$NH_3CHCHCH_2COOH^{2+} + H_2$ (6f)
					$NH_2CCHCH_2C(OH)_2^{2+} + H_2$ (6g)
					$NHCCH_2CH_2C(OH)_2^{2+} + H_2$ (6h)
					$NH_2COCOH(CH_2)_2^{2+} + H_2$ (6i)
					$NH_2CHCH_2CHCOOH^{2+} + H_2$ (6j)
					$NH_2CCH_2COHOCH_2^{2+} + H_2$ (6k)
					$NH_2CHCHCH_2^+ + H_2 + COOH^+$ (6l)
					$NH_2CHCH_2CH_2CO^{2+} + H_2O$ (6m)
					$NH_3 + CH_2CHOCOHCH_2^{2+}$ (6n)

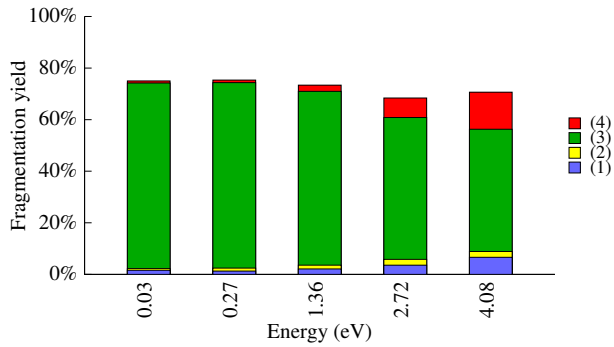


Figure D.7: *Ab initio* molecular dynamics statistics for all 30 trications of GABA. Competition between different Coulomb bond cleavages; fragmentation yield as a function of the excitation energy (in eV). The number in parentheses represents different bonds breaking: (1) $C_{carboxyl} - C_\alpha$, (2) $C_\alpha - C_\beta$, (3) $C_\beta - C_\gamma$ and (4) $C_{carboxyl} - C_\alpha + C_\beta - C_\gamma$ bond cleavages.

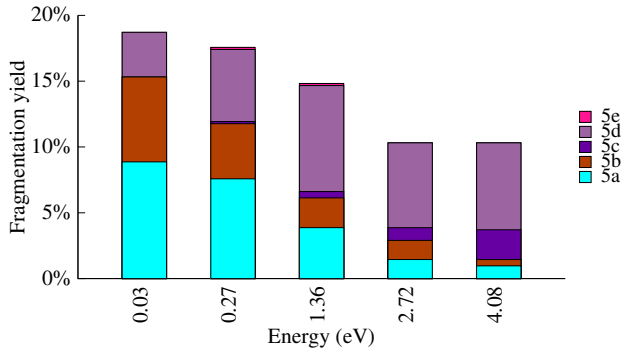


Figure D.8: Statistics of the AIMD simulations showing isomerisation competition for $GABA^{3+}$.

otherwise neutral evaporation is taking place (see Tab. D.2). This criteria allow us to assign double charge to the nitrogen reactive species in coincidence with $COOH^+$ (see 6a-d in Tab. D.2).

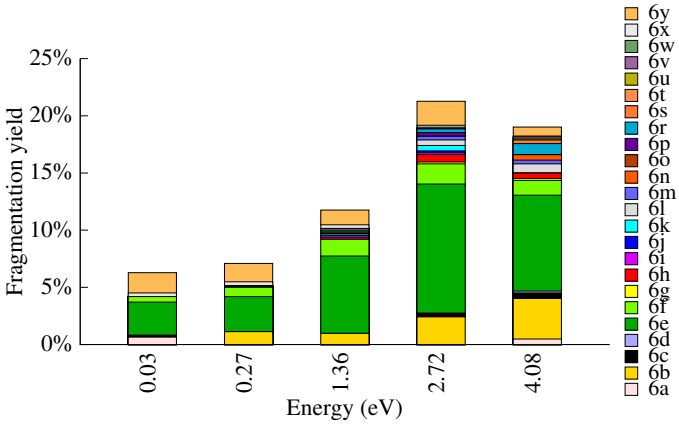


Figure D.9: AIMD simulations showing combined processes competition for GABA^{3+}.

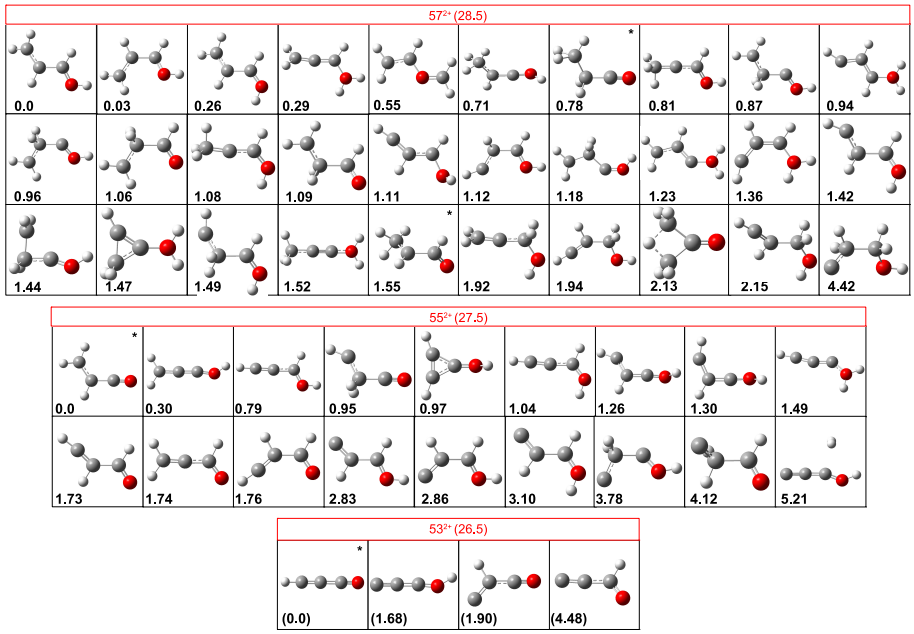


Figure D.10: Possible structures of doubly charged cations at m/z = 26.5, 27.5 and 28.5 amu optimised at the DFT-B3LYP/6-311++G(d,p) level of theory related to ROS - containing oxygen. Relative energies in parentheses between the conformers are in eV. Fragments with star (*) were optimised of the corresponding neutral structure taken from National Institute of Standards and Technology (NIST).

Table D.2: All 44 mechanisms for the fragmentation of $GABA^{3+}$ molecules.

$C_{carboxylic} - C_\alpha$ (1)	$C_\alpha - C_\beta$ (2)	$C_\beta - C_\gamma$ (3)	$C_{carboxylic} - C_\alpha + C_\beta - C_\gamma$ (4)	Isomerisation (5)	Neutral emission and combined processes (6)
$NH_2CH_2CHCH_3^{2+}$ $+COOH^+$ (1a)	$NH_2CHCH_2^{2+}$ $+CH_3COOH^+$ (2a)	$NH_2CH_2^+$ $+CH_2CH_2COOH^{2+}$ (3a)	$NH_2CH_2^+$ $+CH_2CH_2^+$ $+COOH^+$ (4a)	$NH_2CH_2CH_2CH_2COOH^{3+}$ (5a)	$NH_2C(CH_2)_2^{2+}$ $+H_2$ $+COOH^+$ (6a)
$NH_2CH_2CH_2CH_2^{2+}$ $+COOH^+$ (1b)	$NH_2CHCH_3^{2+}$ $+CH_2COOH^+$ (2b)	$NH_2CH_2^+$ $+CH_3CHCOOH^{2+}$ (3b)		$NH_2CHCH_3CH_2COOH^{3+}$ (5b)	$NH_2CHCHCH_2^{2+}$ $+H_2$ $+COOH^+$ (6b)
$NH_2CHCH_2CH_3^{2+}$ $+COOH^+$ (1c)	$NH_3CHCH_2^{2+}$ $+CH_2COOH^+$ (2c)	$NH_2CH_2^+$ $+CH_2OCCH_2OH^{2+}$ (3c)		$NH_3CHCH_2CH_2COOH^{3+}$ (5c)	$NH_2CCH_2CH_2^{2+}$ $+H_2$ $+COOH^+$ (6c)
$NH_3CH_2CHCH_2^{2+}$ $+COOH^+$ (1d)	$NH_2CH_2CH_2^{2+}$ $+CH_2COOH^+$ (2d)	$NH_2CH_2^+$ $+CH_2(CH_2CO)OH^{2+}$ (3d)		$NH_3CH_2CHCH_2COOH^{3+}$ (5d)	$NH_2CCHCH_3^{2+}$ $+H_2^+$ $+COOH$ (6d)
$NH_3CHCHCH_3^{2+}$ $+COOH^+$ (1e)				$NH_2CHCH_2CH_2COH_2O^{3+}$ (5e)	$NH_2CCH_2CH_2COOH^{3+}$ $+H_2$ (6e)
$NH_3CH(CH_2)_2^{2+}$ $+COOH^+$ (1f)					$NH_2CHCHCH_2COOH^{3+}$ $+H_2$ (6f)
					$NH_2CHCH_2CHCOOH^{3+}$ $+H_2$ (6g)
				Neutral emission and combined processes cont. (6)	$CH_2COHOCNH_2CH_2^{3+}$ $+H_2$ (6h)
				$NHCCH_2CH_2COOH^{2+}$ $+H_3^+$ (6r)	$NH_2CCHCH_2C(OH)_2^{3+}$ $+H_2$ (6i)
				$NH_2CHCHCH_3^{2+}$ $+C(OH)_2^+$ (6s)	$NHCCH_2CH_2C(OH)_2^{3+}$ $+H_2$ (6j)
				$NH_2CCHCH_2^+ + COOH^+$ $+H_3^+$ (6t)	$NH_2CCHCH_2C(OH)_2^{3+}$ $+H_2$ (6k)
				$NH_3CH_2CHCH_3^{2+} + CO_2^+$ (6u)	$NH_2CCH_2^{2+} + CH_2COOH^+$ $+H_2$ (6l)
				$NH_2CHCH_2CH_2CO^{2+}$ $+H_2O^+$ (6v)	$NH_2CH^+ + CH_2CH_3^+$ $+COOH^+$ (6m)
				$NH_2CCHCH_2CO^{2+}$ $+H_2O^+ + H_2$ (6w)	$NH_2CHCH_2^{2+}$ $+OCCH_3^+ + OH$ (6n)
				$NH_2CHCH_2CH_2^+ + H^+$ $+COOH^+$ (6x)	$NH_2CHCH_2^+$ $+CH_2COH_2O^{2+} + OH$ (6o)
				$NH_2CCH_3^+ + CH_3^+$ $+COOH^+$ (6y)	$NH_2CH_2^+$ $+CH_2COH^+ + CH_2O$ (6p)

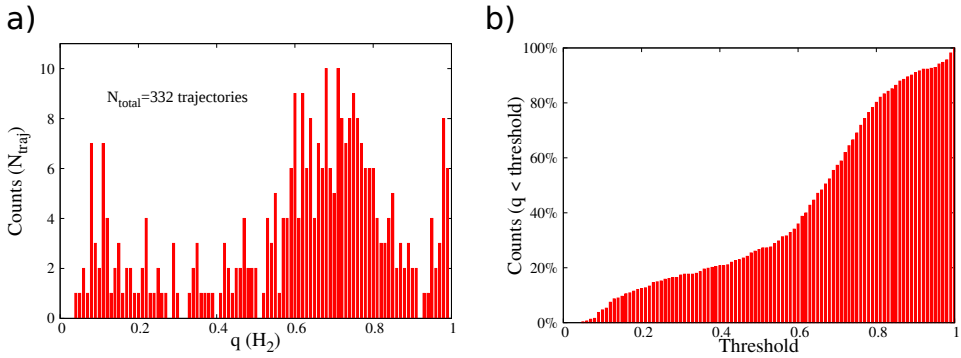


Figure D.11: The counts of the 323 fragmentation mechanisms related with H_2 emission observed in combined processes. Panel a) shows the counts for different charges of the hydrogen molecules (q H_2); panel b) shows the normalised (to 100%) counts for q<threshold as a function of the threshold.

Appendix E

Appendices
Additional information for N-acetylglycine$^+$

A detailed study of the potential energy surface (PES) of neutral and singly positively charged N-acetylglycine is presented here. In the following figures we present the critical points (minima and transition states) corresponding to:

- Structure of neutral conformers of N-acetylglycine (Fig. E.1);
- Isomerisation pathways for neutral N-acetylglycine conformers (Fig. E.2);
- Structure of singly charged conformers of N-acetylglycine (Fig. E.3);
- Isomerisation pathways for singly charged N-acetylglycine conformers (Fig. E.4);
- Structure of singly charged diols of N-acetylglycine (Fig. E.5);
- Isomerisation pathways for singly charged N-acetylglycine diols (Fig. E.6).

Figure E.1: Structure of neutral conformers of N-acetylglycine; "**c**" (*cis*) and "**t**" (*trans*) notation underline the position of the atoms with respect to peptide bond; star "*****" indicates the rotamers. Relative energies are given in kcal mol^{-1} at the DFT-B3LYP/6-311++G(d,p) level of theory with respect to the most stable isomer of the neutral molecule (t6) including the zero point energy correction.

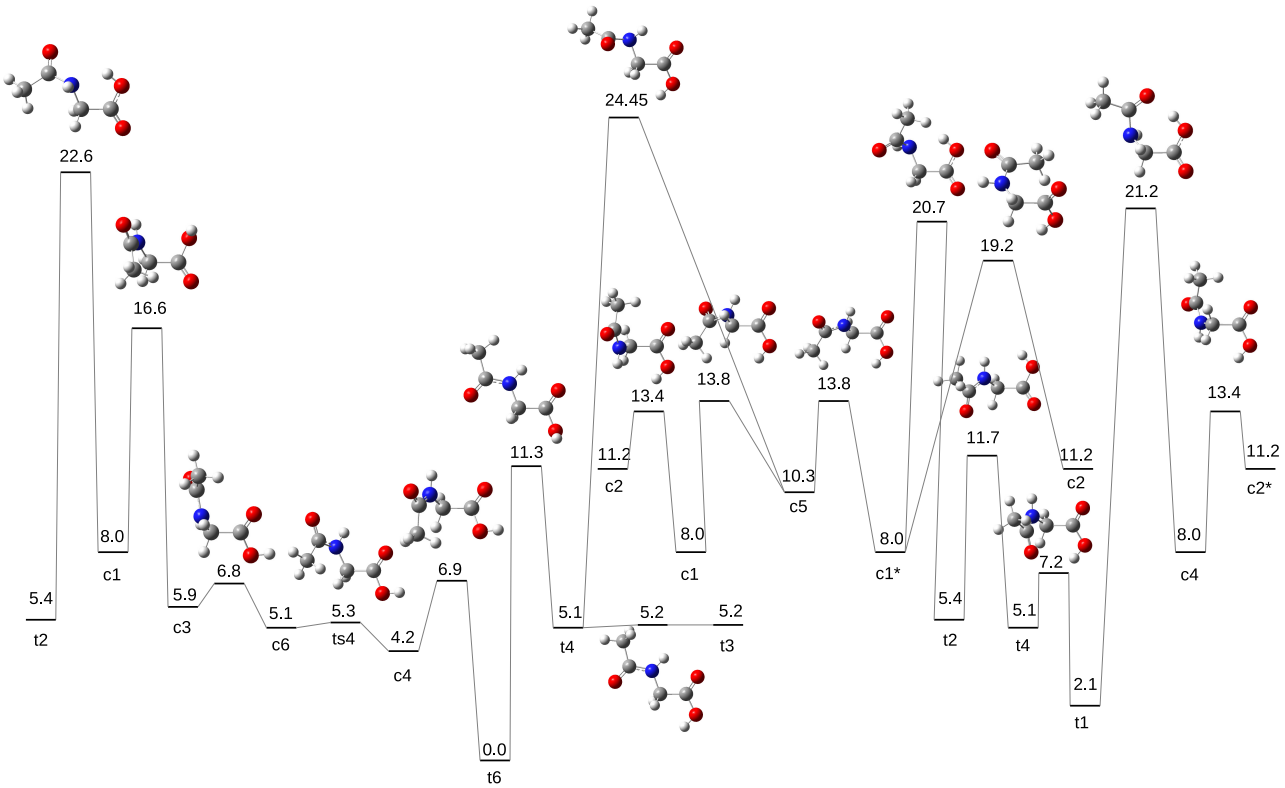


Figure E.2: Isomerisation pathways for neutral *N*-acetylglycine. Relative energies are given in kcal mol^{-1} with respect to the most stable neutral conformer of *N*-acetylglycine (t6) at DFT-B3LYP/6-311++G(d,p) level of theory including the zero point energy correction.

Figure E.3: Structure of singly charged conformers of N-acetylglycine; "**c**" (*cis*) and "**t**" (*trans*) notation underline the position of the atoms with respect to peptide bond; star "*****" indicates the rotamers. Relative energies are given in kcal mol^{-1} with respect to the t6+ conformer at the DFT-B3LYP/6-311++G(d,p) level of theory including the zero point energy correction.

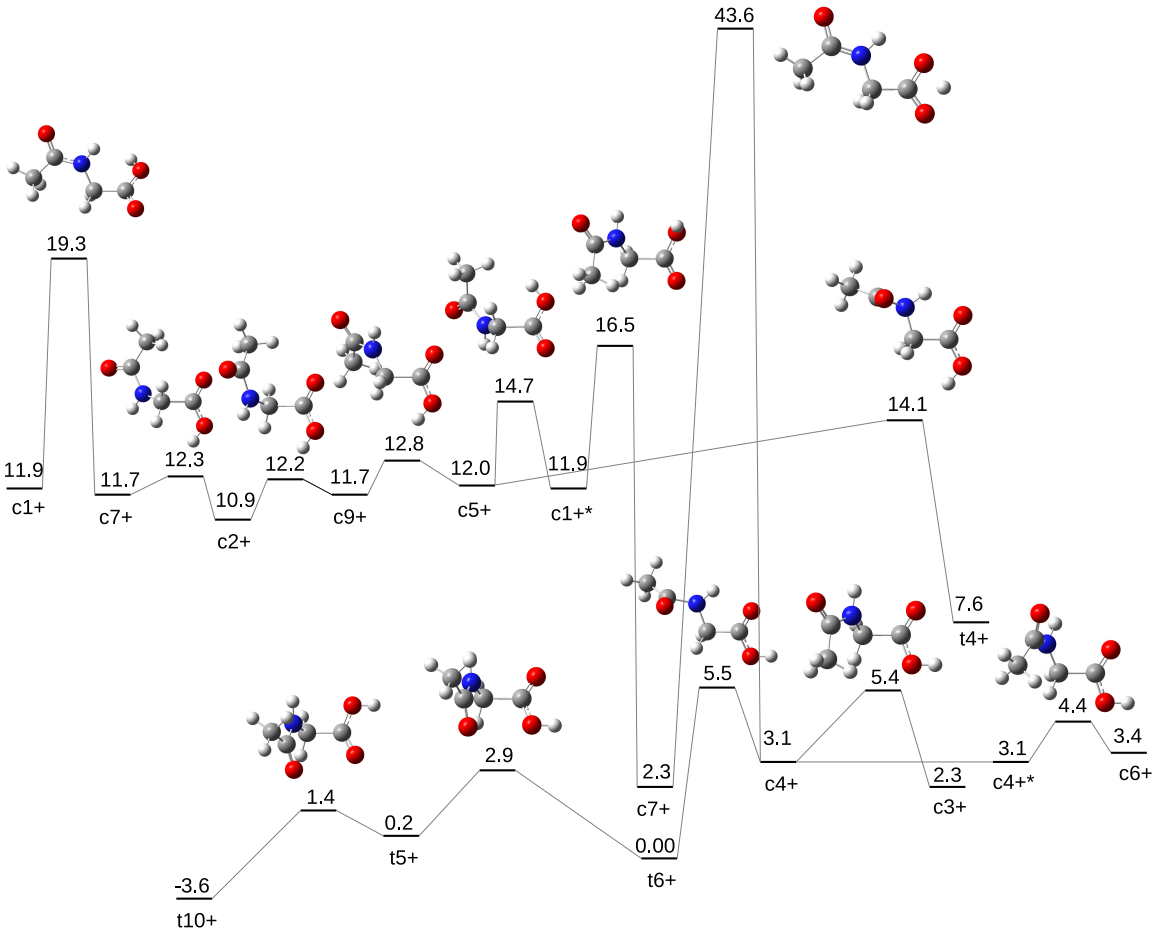


Figure E.4: Isomerisation pathways for singly charged *N*-acetylglycine. Relative energies are given in kcal mol^{-1} with respect to the t6+ conformer of *N*-acetylglycine at DFT-B3LYP/6-311++G(d,p) level of theory including the zero point energy correction.

Figure E.5: Structure of singly charged diols (d) of N-acetylglycine; "**c**" (*cis*) and "**t**" (*trans*) notation underline the position of the atoms with respect to peptide bond; apostrophe "'" indicates the different relative position of two hydrogens in a geminal diol cations. Relative energies are given in kcal mol^{-1} with respect to the t6+ conformer at the DFT-B3LYP/6-311++G(d,p) level of theory including the zero point energy correction.

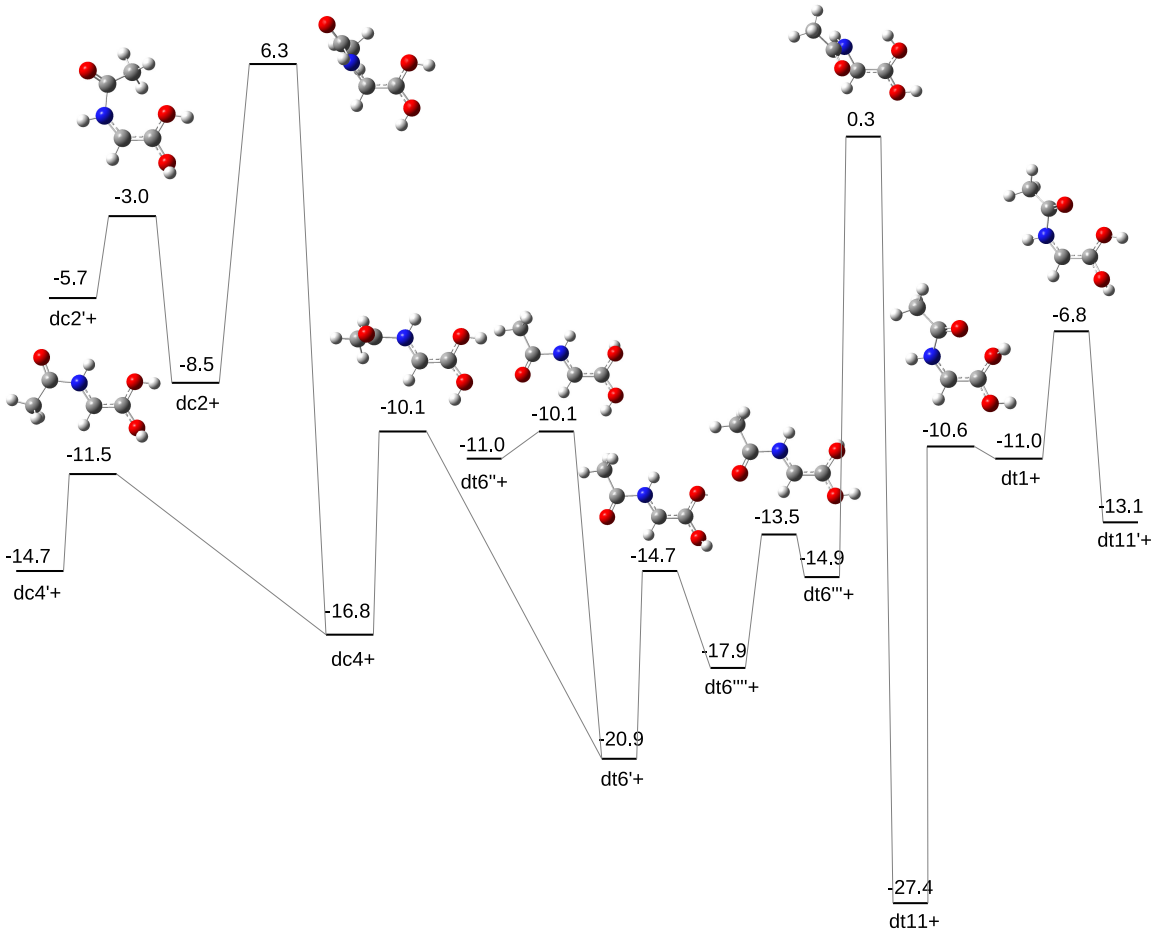


Figure E.6: Isomerisation pathways for singly charged *N*-acetylglycine diols; apostrophe " ' " indicates the different relative position of two hydrogens in a geminal diol cations. Relative energies are given in kcal mol^{-1} with respect to the t6+ conformer of *N*-acetylglycine at DFT-B3LYP/6-311++G(d,p) level of theory including the zero point energy correction.

Appendix F

Appendices
Additional information for neutral clusters of β-alanine

F.1 Thermostats

The input files to run molecular dynamic (MD) simulations with AMBER[136–139] were prepared using antechamber tools.[393] General Amber force field[147] were used with RESP charges[148] parametrised at the HF/6-31G* level of theory with the Gaussian09 program.[195] Energy minimisation was carried out to remove the possible non-physical, accidental contacts after parameterising. Then, different thermostats were tested in order to obtain physically meaningful results. For each thermostat the different parameters were converged at each stage of the MD simulation. First, the system was heated from T=0K to the target temperature and then propagated at this temperature. Three different thermostats were checked: Berendsen,[346] Andersen[394] and Langevin.[347] The best results were obtained using the Berendsen thermostat for the heating stage and Langevin thermostat for the production dynamics. However, for both thermostats the results were unsatisfying since we observed high oscillations in energy and temperature and a further extensive benchmark study for thermal equilibration was performed.

Berendsen thermostat convergence

Firstly, the conditions for the Berendsen thermostat were optimised. We performed an extensive study to choose the best set of parameters: time step,Δt (fs), and temperature coupling parameter, taut (ps), by using the Berendsen thermostat for all clusters sizes consider in the manuscript. To this, during the heating stage from T=0K to T=(50, 100, 150, 200, 250, 300)K we checked all combination for taut=(0.5, 1, 1.5, 2, 2.5, 3, 5, 10, 15, 20, 25, 50, 100, 150, 200 and 250 picoseconds) and Δt=(0.0001, 0.0002, 0.0003, 0.0004, 0.0005, 0.0006, 0.0007, 0.0008, 0.0009 and 0.001 fs); over 5000 simulations. After analysing the simulations we concluded that taut=10 ps and Δt=0.2 fs are the most appropriate parameters, since the

temperature is well equilibrated for each cluster size and for each excitation energy during the propagation stage. The plots for each cluster size and temperature with the chosen parameters are presented in tables F.1-F.4.

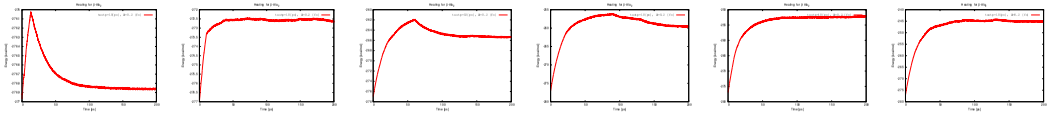
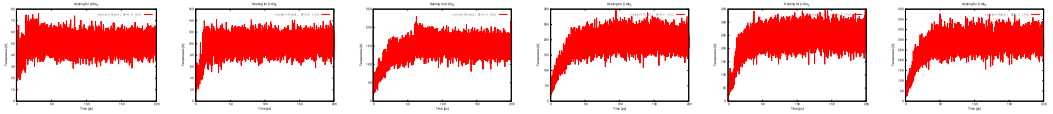
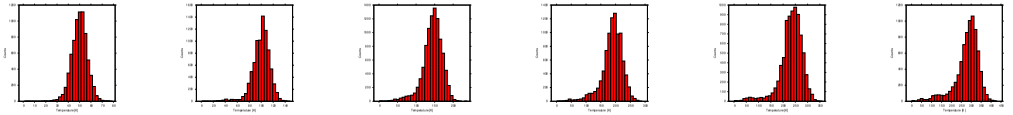
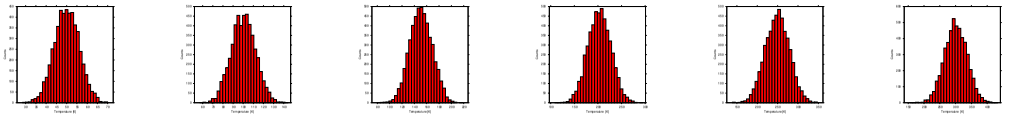
Quantity		Temperature					
		50K	100K	150K	200K	250K	300K
Parameter sets	taut[ps]	10	10	10	10	10	10
	Δt[fs]	0.2	0.2	0.2	0.2	0.2	0.2
Energy curves							
Temperature curves							
Histogram full simulation							
Histogram last 100 ps							

Table F.1: Parametrisation using Berendsen thermostat for (β-Ala)$_2$

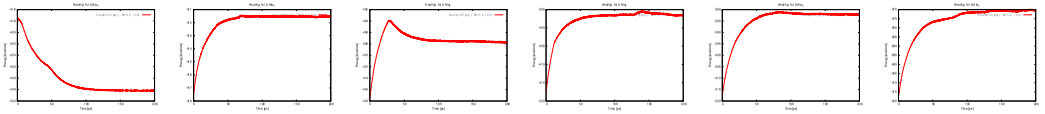
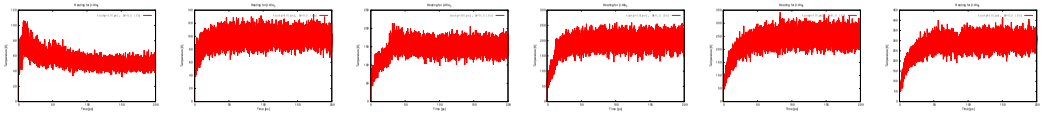
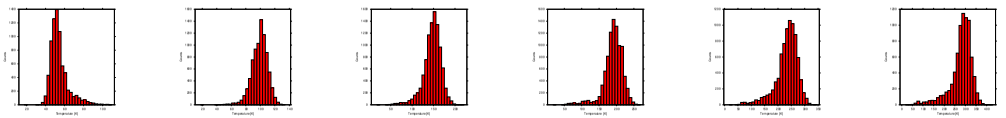
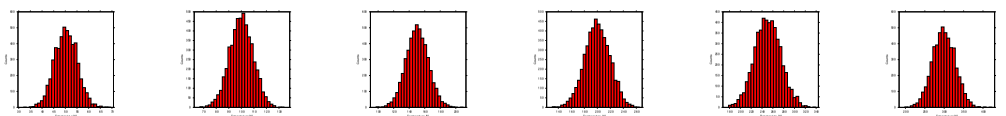
Quantity		Temperature					
		50K	100K	150K	200K	250K	300K
Parameter sets	tautp[ps]	10	10	10	10	10	10
	Δt[fs]	0.2	0.2	0.2	0.2	0.2	0.2
Energy curves							
Temperature curves							
Histogram full simulation							
Histogram last 100 ps							

Table F.2: Parametrisation using Berendsen thermostat for $(\beta\text{-Ala})_3$

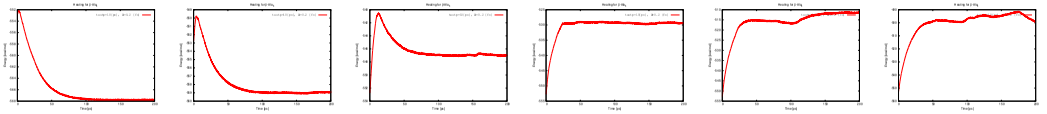
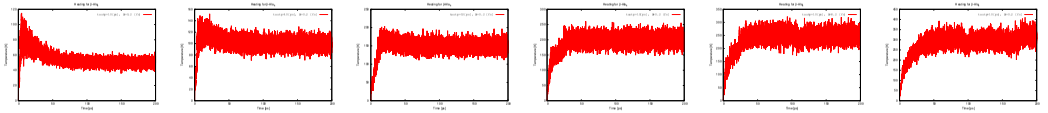
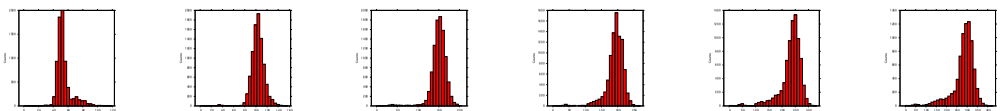
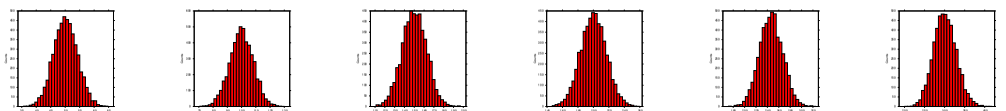
Quantity		Temperature					
		50K	100K	150K	200K	250K	300K
Parameter sets	tautp[ps]	10	10	10	10	10	10
	Δt[fs]	0.2	0.2	0.2	0.2	0.2	0.2
Energy curves							
Temperature curves							
Histogram full simulation							
Histogram last 100 ps							

Table F.3: Parametrisation using Berendsen thermostat for (β-Ala)$_4$

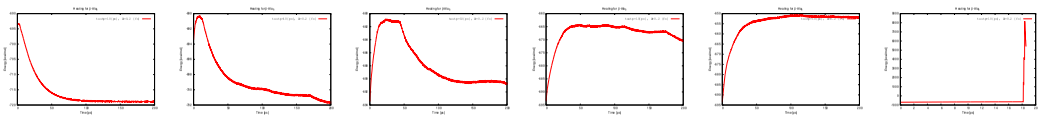
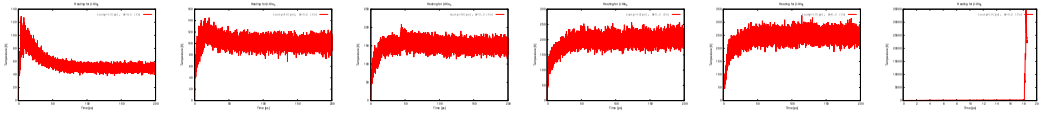
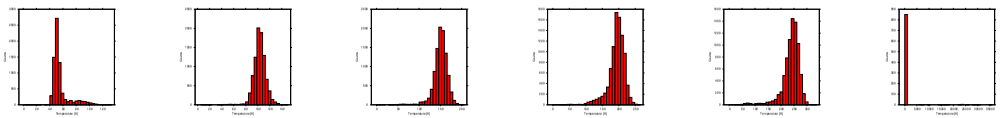
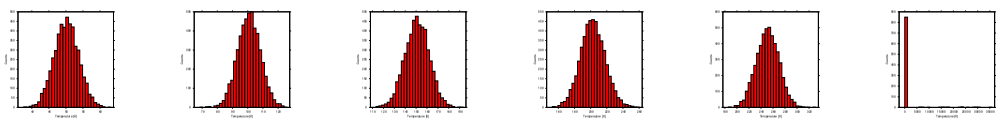
Quantity		Temperature					
		50K	100K	150K	200K	250K	300K
Parameter sets	tautp[ps]	10	10	10	10	10	10
	Δt[fs]	0.2	0.2	0.2	0.2	0.2	0.2
Energy curves							
Temperature curves							
Histogram full simulation							
Histogram last 100 ps							

Table F.4: Parametrisation using Berendsen thermostat for (β-Ala)$_5$ For 300K cluster is not stable, it thermally breaks.

Langevin thermostat convergence

Secondly, we optimised the conditions for the Langevin thermostat. We performed an extensive study to choose the best set of parameters: time step, Δt (fs), and the collision frequency, γ (ps^{-1}). For the Langevin thermostat we considered the same cluster sizes and temperature values as for the Berendsen thermostat. Simulations were performed with γ=(1, 2, 3, 4, 5, 10, 20, 25, 50, 60, 70, 80, 90, 100 ps^{-1}) and Δt=(0.1, 0.2, ... 1.5, 2.0 fs). After these simulations we concluded that γ=5 ps^{-1} (weak coupling to approximate the micro-canonical ensemble, NVE) and Δt=0.2 fs are the most appropriate parameters, where the energy and the temperature are well equilibrated for each cluster size and for each temperature during the propagation stage. The plots for each cluster size and temperature with the chosen parameters are presented in Tables F.5-F.8.

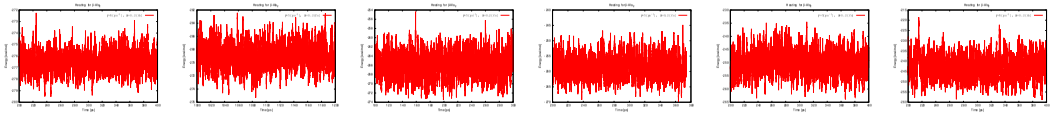
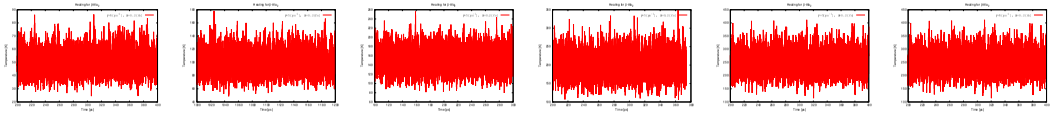
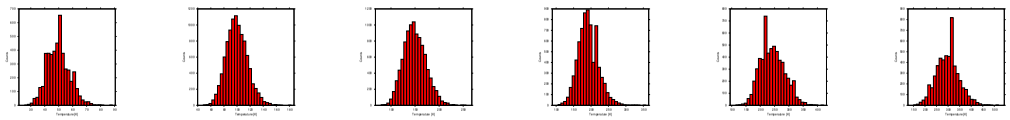
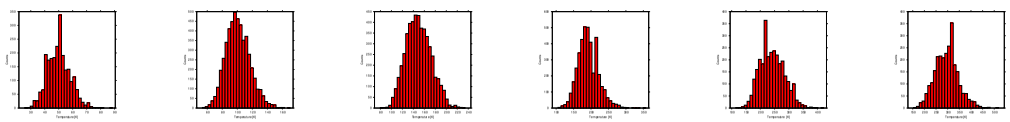
Quantity	Temperature					
	50K	100K	150K	200K	250K	300K
Energy curves						
Temperature curves						
Histogram full simulation						
Histogram last 100 ps						

Table F.5: Parametrisation using Langevin thermostat for $(\beta\text{-Ala})_2$. Parameters for each temperature are: γ=5 ps^{-1} and Δt=0.2 fs.

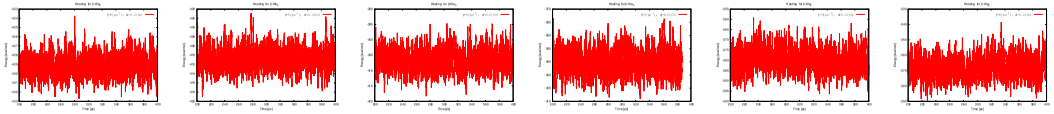
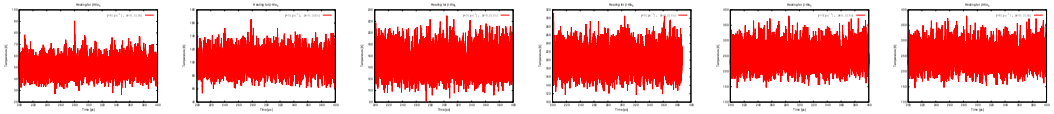
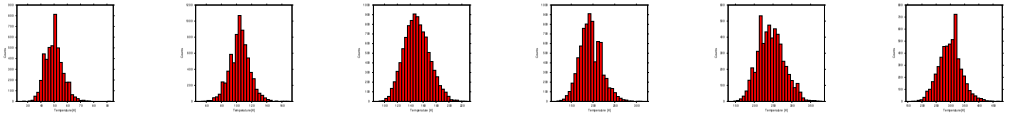
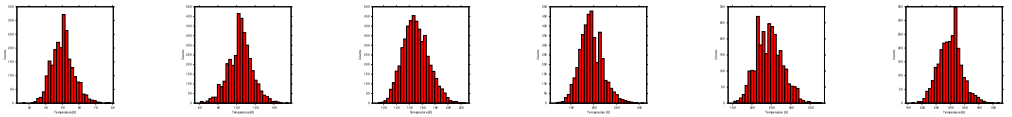
Quantity	Temperature					
	50K	100K	150K	200K	250K	300K
Energy curves						
Temperature curves						
Histogram full simulation						
Histogram last 100 ps						

Table F.6: Parametrisation using Langevin thermostat for (β-Ala)$_3$. Parameters for each temperature are: γ=5 ps^{-1} and Δt=0.2 fs.

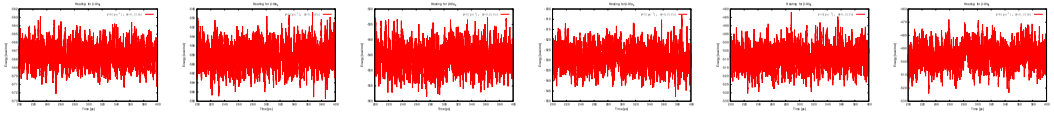
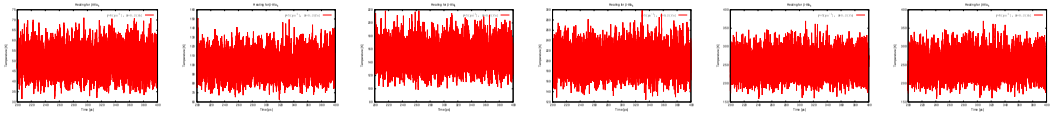
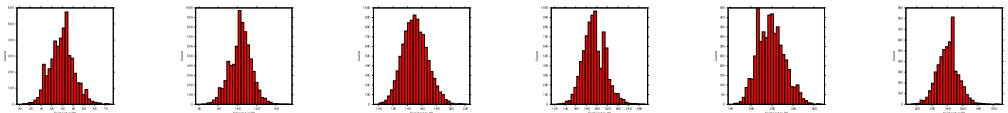
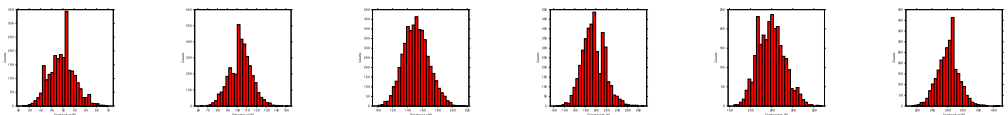
Quantity	Temperature					
	50K	100K	150K	200K	250K	300K
Energy curves						
Temperature curves						
Histogram full simulation						
Histogram last 100 ps						

Table F.7: Parametrisation using Langevin thermostat for $(\beta\text{-Ala})_4$. Parameters for each temperature are: γ=5 ps^{-1} and Δt=0.2 fs.

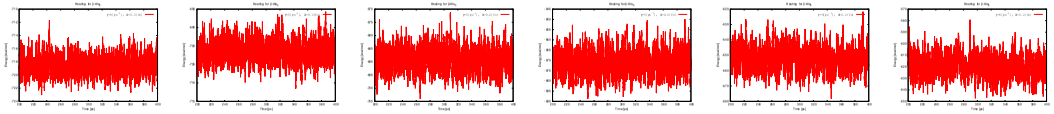
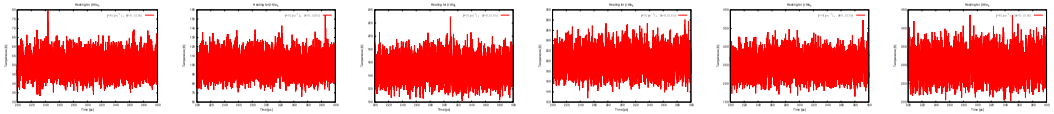
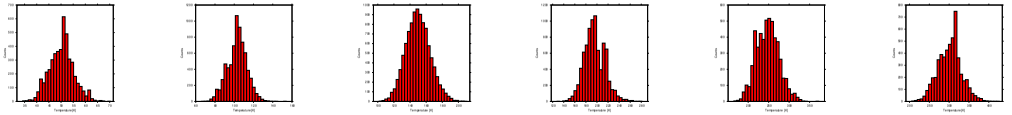
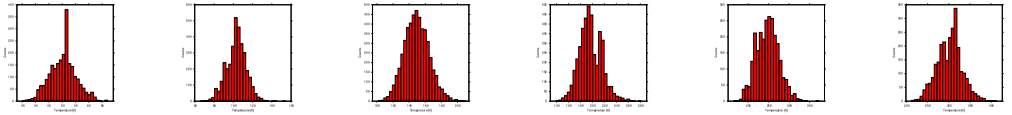
Quantity	Temperature					
	50K	100K	150K	200K	250K	300K
Energy curves						
Temperature curves						
Histogram full simulation						
Histogram last 100 ps						

Table F.8: Parametrisation using Langevin thermostat for (β-Ala)$_5$. Parameters for each temperature are: γ=5 ps^{-1} and Δt=0.2 fs.

Summary of the thermostats convergence

A large set of various parameters to run classical molecular dynamics simulations for both processes heating and propagation were considered. After the evaluation of these parameters we obtained the most stable ones for each temperature at the heating and production stages, using the Berendsen and Langevin thermostats, respectively. We checked that the total energy and temperature do not fluctuate as a function of time to be sure that system is well equilibrated. Some slight deviations in the temperature are observed due to the hydrogen bonds creation and annihilation, and due to van der Waals interactions.

F.2 Benchmarking

The benchmarking study has been carried out evaluating geometries, relative energies and bonding properties. Table F.9 presents differences in the geometries obtained with the MP2 method and with each of the considered functionals: B3LYP, dispersion corrected B97D, M06 and its variation M06-2X and MPWB1K in combination with the 6-311++G(d,p) basis set. Benchmarking of some electron properties in the bond critical points (BCPs) of H bonds for the most stable dimer (2.1) are presented in Tab.F.10. Fig.F.1 presents a general comparison of the relative energies of 19 conformers of dimers after geometry optimisation with the MP2 method and all considered DFT functionals. Single point energies at the CCSD(T)/6-311++G(d,p) level of theory were obtained from geometries optimised at MP2 level of theory for the most stable 11 conformers. Fig.F.2 shows benchmarking of the relative energies computed with the DFT functionals and their correlation with the MP2 and CCSD(T) methods. Almost complete energetic similarity measured with the Pearson coefficient is found for the M06-2X functional in comparison with MP2 or CCSD(T).

Table F.9: Root mean square deviations, $RMSD$ in Å, for the structures in $(\beta - ala)_2$ calculated for each considered functional with respect to the MP2 results. The biggest and the smallest values of $RMSD$ are highlighted in bold.

Conformer	B3LYP	B97D	M06	M06-2X	MPWB1K
2.1	0.08	**0.04**	0.04	0.05	0.05
2.2	0.13	0.11	0.17	0.13	0.13
2.3	0.10	0.09	0.12	0.13	0.13
2.4	0.14	0.11	0.25	0.29	0.25
2.5	0.28	0.26	0.31	**0.32**	0.31
2.6	0.12	0.12	0.35	0.25	0.30
2.7	0.19	0.18	0.34	0.14	0.25
2.8	0.08	0.12	0.12	0.10	0.08
2.9	**0.87**	**0.73**	**0.82**	0.16	**0.85**
2.10	0.16	0.05	0.09	0.06	0.07
2.11	0.45	0.05	0.17	0.09	0.08
2.12	0.54	0.10	0.13	0.12	0.13
2.13	0.56	0.11	0.08	0.06	0.14
2.14	0.07	0.05	**0.03**	0.04	**0.04**
2.15	0.22	0.16	0.11	0.17	0.25
2.16	0.10	0.09	0.08	0.17	0.09
2.17	**0.06**	0.05	**0.03**	**0.03**	**0.04**
2.18	0.70	0.11	0.19	0.10	0.10
2.19	0.44	0.21	0.06	0.22	0.46
Average	0.28	0.14	0.18	0.14	0.20

Table F.10: Electron density and its Laplacian computed for the most stable conformer of $(\beta - ala)_2$ at the bond critical points BCP1 and BCP2 (see bonds labels in the main text).

	MP2	B3LYP	B97D	M06	M06-2X	MPWB1K
ρ (BCP1)	0.050	0.050	0.052	0.049	0.050	0.049
ρ (BCP2)	0.050	0.050	0.052	0.049	0.050	0.049
$\nabla^2\rho$ (BCP1)	0.100	0.100	0.093	0.108	0.100	0.104
$\nabla^2\rho$ (BCP2)	0.100	0.100	0.093	0.108	0.100	0.104

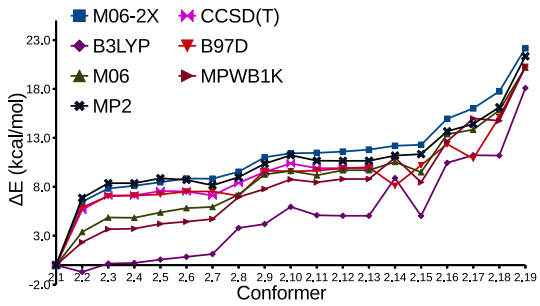


Figure F.1: Benchmarking of 19 conformers of $(\beta - ala)_2$: relative stability between all conformers at the different levels of theory that we considered; Relative energies are given in kcal mol $^{-1}$ with respect to the most stable conformer.

F.3 Isomers

Using the computational strategy described in the main manuscript and employing the M06-2X functional with the 6-311++G(d,p) basis set we evaluated the relative stability of numerous isomers (m) of different cluster sizes (n). We use the nomenclature (n,m), e.g. conformer 3.9 is the ninth conformer of a cluster with three beta-alanine residues. A complete study of the isomers of neutral β-alanine molecules is presented in Figures F.3-F.10. We present the optimised geometries, stabilised mainly through hydrogen bonds, and their relative energies. The monomers were obtained after reoptimisation of geometries taken from a previous study.[55]

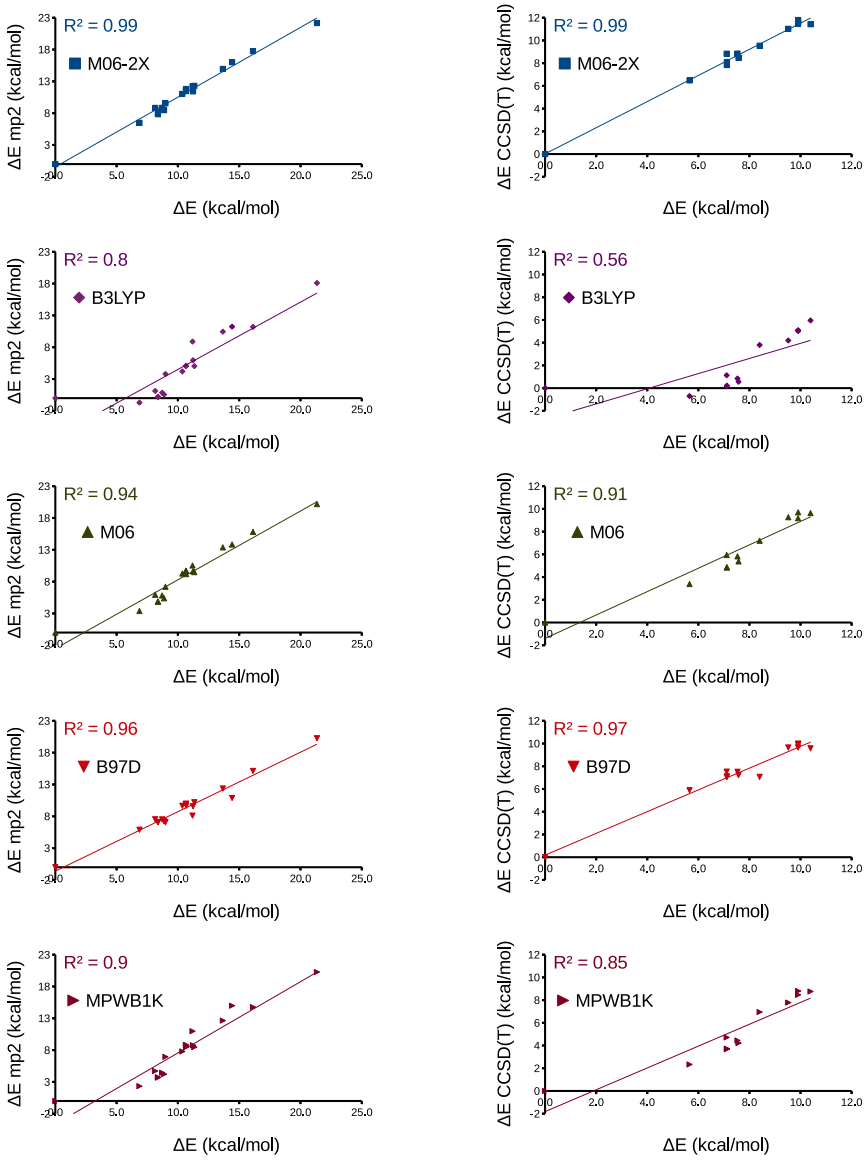


Figure F.2: Benchmarking of 19 and 11 conformers of $(\beta - ala)_2$: correlations of the different functionals with the MP2 and CCSD(T) results, respectively. R is the Pearson correlation coefficient. Relative energies are given in kcal mol $^{-1}$ with respect to the most stable conformer (see all structures in the Fig.F.4).

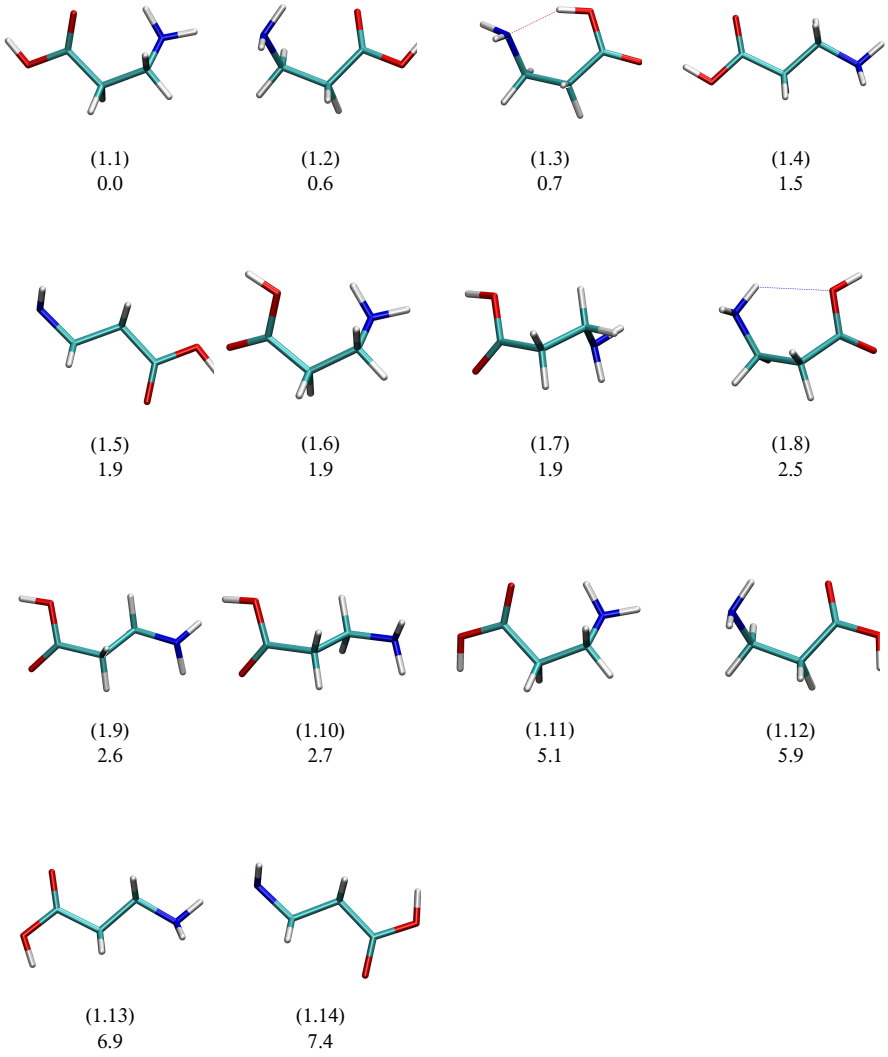


Figure F.3: optimised geometries of neutral monomers (1.1-1.14) of β-alanine at the DFT-M06-2X/6-311++G(d,p) level of theory. Relative energies with respect to the most stable neutral monomer (1.1) in kcal mol^{-1}.

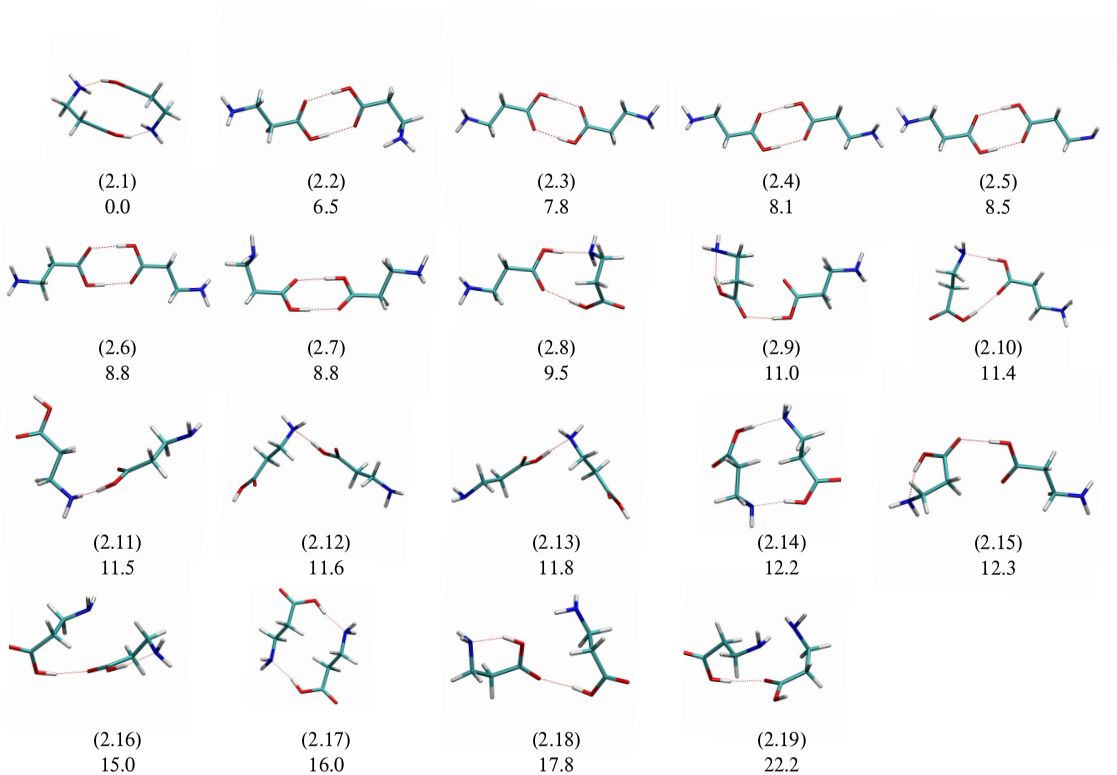


Figure F.4: optimised geometries of neutral dimers (2.1-2.19) of β-alanine at the DFT-M06-2X/6-311++G(d,p) level of theory. Relative energies with respect to the most stable neutral dimer (2.1) in kcal mol^{-1}. Dashed lines shows hydrogen (H) bonds with donor (D)-acceptor (A) atoms with the cut-off distance $\leq$ 3.2 and D-H-A angle $\leq$ 50 degree.[355] The colours of hydrogen bond: red, blue and turquoise underline the donor atom: oxygen, nitrogen and carbon, respectively.

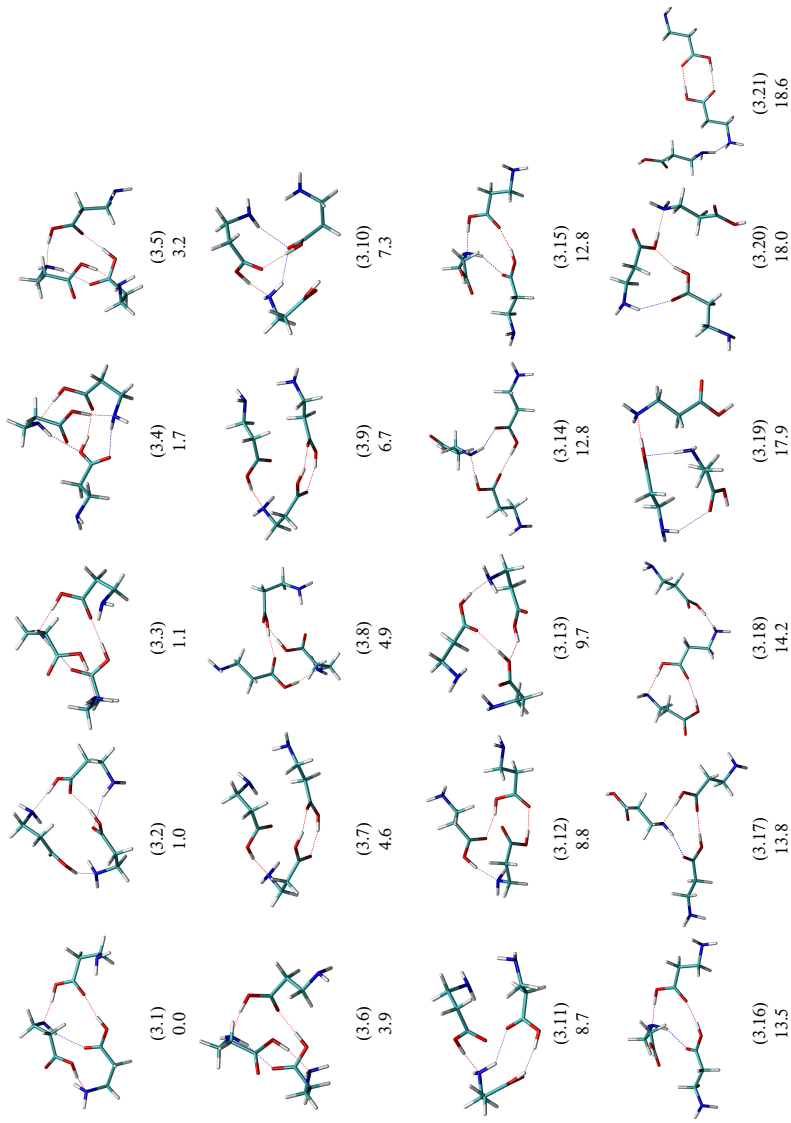


Figure F.5: optimised geometries of neutral trimers (3.1-3.21) of β-alanine at the DFT-M06-2X/6-311++G(d,p) level of theory. Relative energies with respect to the most stable neutral trimer (3.1) in kcal mol^{-1}. Dashed lines shows hydrogen (H) bonds with donor (D)-acceptor (A) atoms with the cut-off distance $\leq$ 3.2 and D-H-A angle $\leq$ 50 degree.[355] The colours of hydrogen bond: red, blue and turquoise underline the donor atom: oxygen, nitrogen and carbon, respectively.

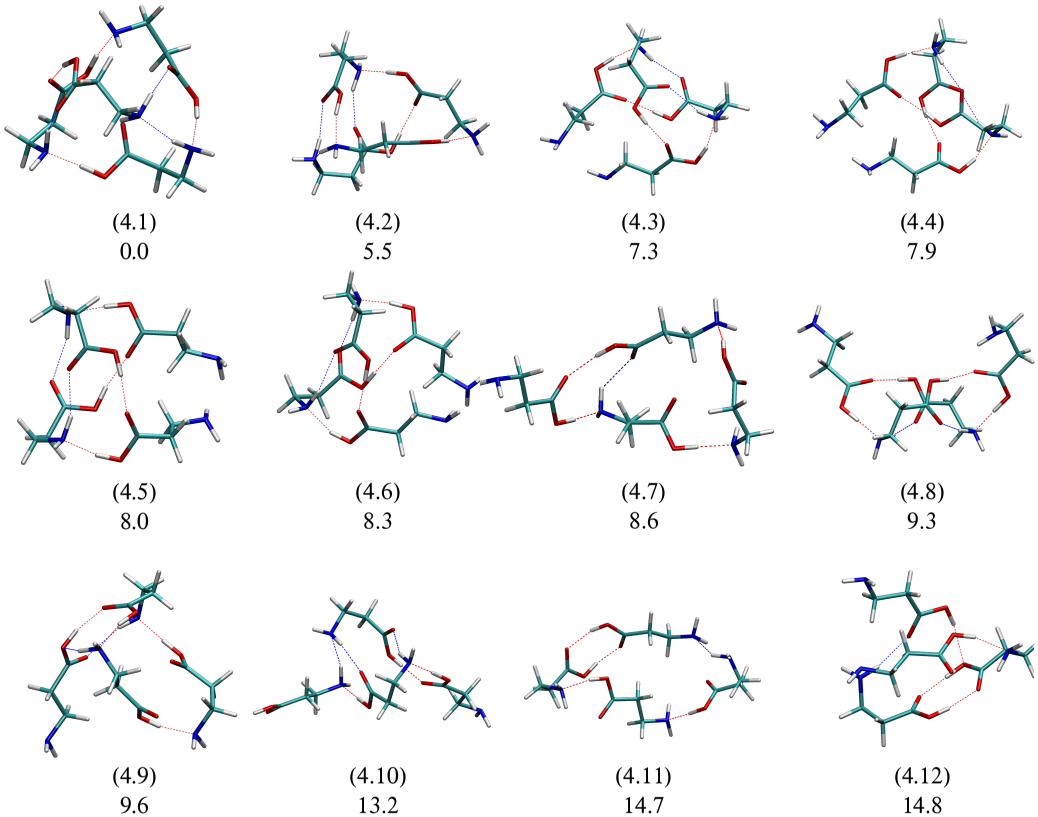


Figure F.6: optimised geometries of neutral tetramers (4.1-4.12) of β-alanine at the DFT-M06-2X/6-311++G(d,p) level of theory. Relative energies with respect to the most stable neutral tetramer (4.1) in kcal mol^{-1}. Dashed lines shows hydrogen (H) bonds with donor (D)-acceptor (A) atoms with the cut-off distance $\leq$ 3.2 and D-H-A angle $\leq$ 50 degree.[355] The colours of hydrogen bond: red, blue and turquoise underline the donor atom: oxygen, nitrogen and carbon, respectively.

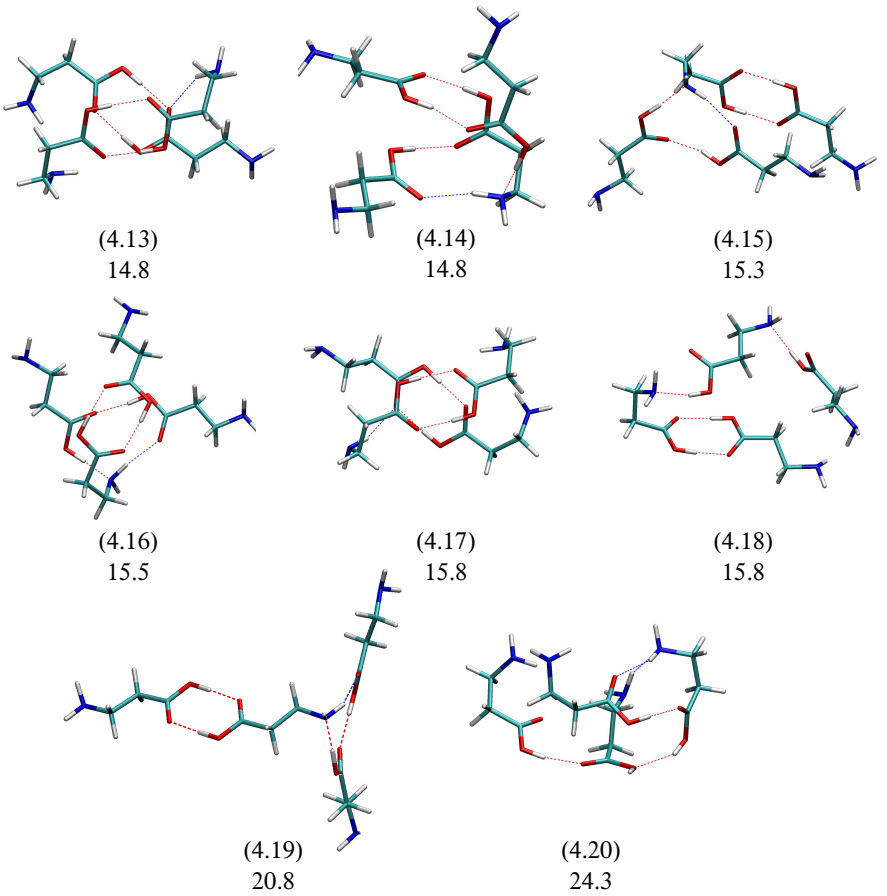


Figure F.7: optimised geometries of neutral tetramers (4.13-4.20) of β-alanine at the DFT-M06-2X/6-311++G(d,p) level of theory. Relative energies with respect to the most stable neutral tetramer (4.1) in kcal mol^{-1}. Dashed lines shows hydrogen (H) bonds with donor (D)-acceptor (A) atoms with the cut-off distance $\leq$ 3.2 and D-H-A angle $\leq$ 50 degree.[355] The colours of hydrogen bond: red, blue and turquoise underline the donor atom: oxygen, nitrogen and carbon, respectively.

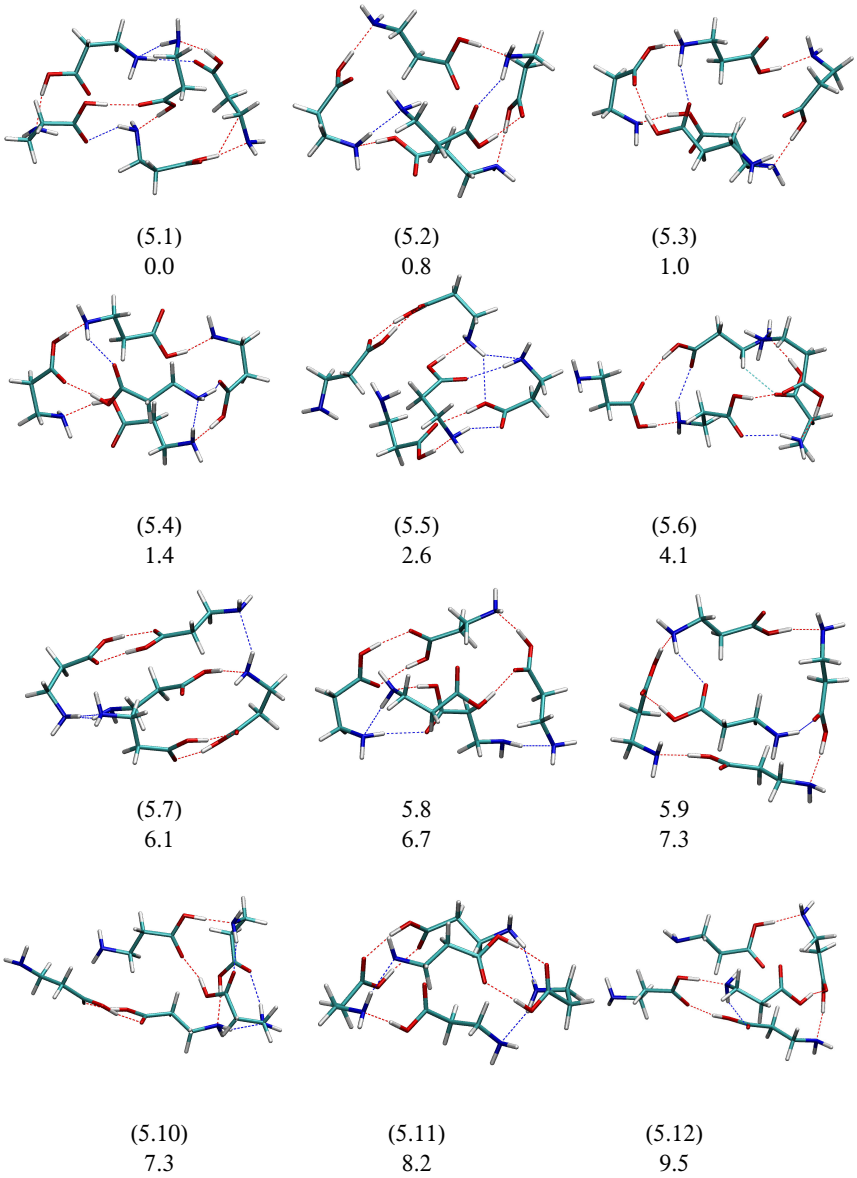


Figure F.8: Optimised geometries of neutral pentamers (5.1-5.12) of β-alanine at the DFT-M06-2X/6-311++G(d,p) level of theory. Relative energies with respect to the most stable neutral tetramer (5.1) in kcal mol^{-1}. Dashed lines shows hydrogen (H) bonds with donor (D)-acceptor (A) atoms with the cut-off distance $\leq$ 3.2 and D-H-A angle $\leq$ 50 degree.[355] The colours of hydrogen bond: red, blue and turquoise underline the donor atom: oxygen, nitrogen and carbon, respectively.

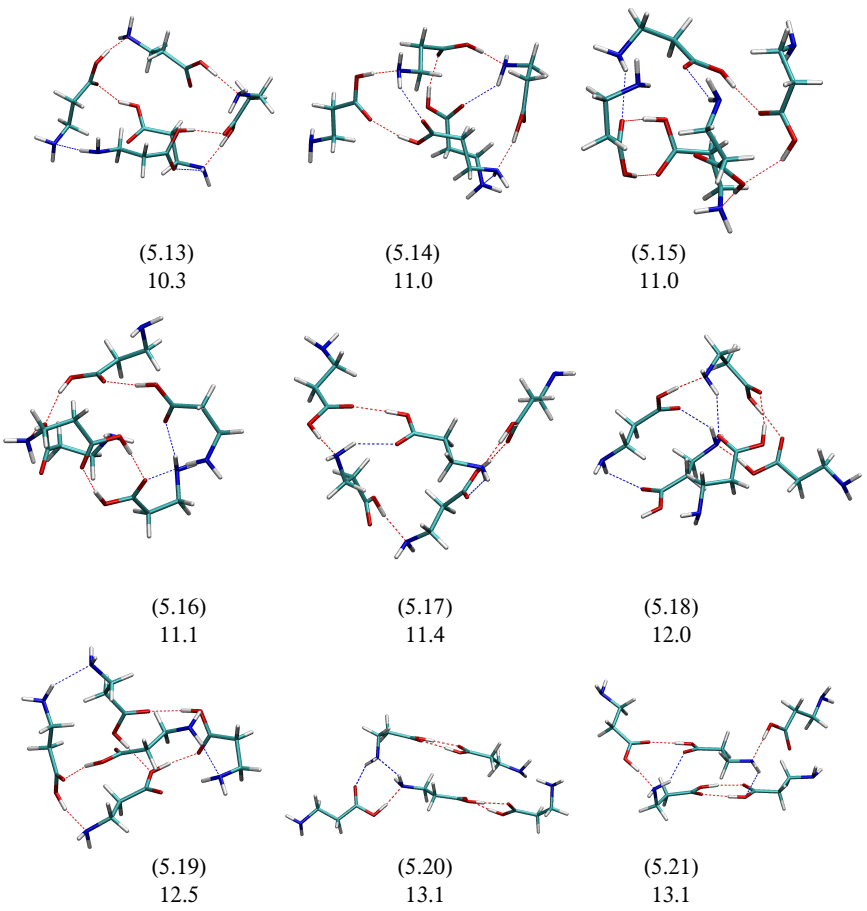


Figure F.9: Optimised geometries of neutral pentamers (5.13-5.21) of β-alanine at the DFT-M06-2X/6-311++G(d,p) level of theory. Relative energies with respect to the most stable neutral tetramer (5.1) in kcal mol^{-1}. Dashed lines shows hydrogen (H) bonds with donor (D)-acceptor (A) atoms with the cut-off distance $\leq$ 3.2 and D-H-A angle $\leq$ 50 degree.[355] The colours of hydrogen bond: red, blue and turquoise underline the donor atom: oxygen, nitrogen and carbon, respectively.

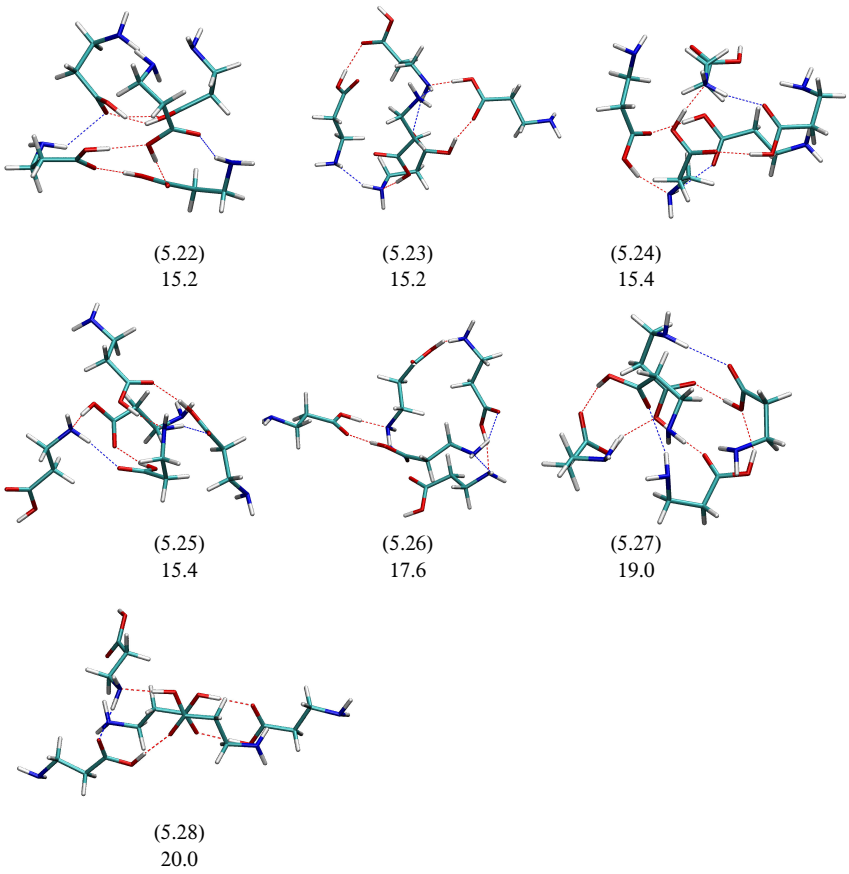


Figure F.10: Optimised geometries of neutral pentamers (5.22-5.28) of β-alanine at the DFT-M06-2X/6-311++G(d,p) level of theory. Relative energies with respect to the most stable neutral tetramer (5.1) in kcal mol^{-1}. Dashed lines shows hydrogen (H) bonds with donor (D)-acceptor (A) atoms with the cut-off distance $\leq$ 3.2 and D-H-A angle $\leq$ 50 degree.[355] The colours of hydrogen bond: red, blue and turquoise underline the donor atom: oxygen, nitrogen and carbon, respectively.

Finally, we present a thorough study based on wave function analysis techniques for the most stable conformer of each cluster size. We focused on the estimated stabilisation energies for all evaluated hydrogen bonds including a detailed analysis of the type of the donor hydrogen atom, acceptor hydrogen atom, acceptor and donor orbital energy differences and Fock-diagonal NBO Fock matrix elements.

Table F.11: Estimated stabilisation energy ($\Delta E_{ij}^{(2)}$) (in kcal mol^{-1}) of all the interaction between atoms involved in the hydrogen bonds (HBs) for the most stable dimer. Acceptor and donor orbital energy differences ($\epsilon_j^{(NL)} - \epsilon_i^{(L)}$) in a.u. and off-diagonal NBO Fock matrix elements (F_{ij}) in a.u. are also presented.

DIMER (β-ala)$_2$

HBs index	Donor NBO (i)					Acceptor NBO (j)					$\Delta E_{ij}^{(2)}$	$\epsilon_j^{(NL)} - \epsilon_i^{(L)}$	F_{ij}
1	BD	C	2	N	3	BD*	O	18	H	22	0.28	1.25	0.017
	BD	N	3	H	7	BD*	O	18	H	22	0.28	1.15	0.016
	BD	N	3	H	8	BD*	O	18	H	22	1.62	1.15	0.039
	CR	N	3	-	-	BD*	O	18	H	22	0.38	15.02	0.069
	LP	N	3	-	-	BD*	O	18	H	22	34.94	0.86	0.155
	BD	O	18	H	22	RY*	N	3	-	-	0.07	1.56	0.01
	BD	O	18	H	22	BD*	C	2	N	3	0.07	1.22	0.008
	BD	O	18	H	22	BD*	N	3	H	7	0.15	1.32	0.013
	BD	O	18	H	22	BD*	N	3	H	8	0.06	1.33	0.008
2	BD	O	5	H	9	RY*	N	16	-	-	0.07	1.56	0.01
	BD	O	5	H	9	BD*	C	15	N	16	0.07	1.22	0.008
	BD	O	5	H	9	BD*	N	16	H	20	0.15	1.32	0.013
	BD	O	5	H	9	BD*	N	16	H	21	0.06	1.33	0.008
	BD	C	15	N	16	BD*	O	5	H	9	0.27	1.25	0.017
	BD	N	16	H	20	BD*	O	5	H	9	0.28	1.15	0.016
	BD	N	16	H	21	BD*	O	5	H	9	1.62	1.15	0.039
	CR	N	16	-	-	BD*	O	5	H	9	0.38	15.02	0.069
	LP	N	16	-	-	BD*	O	5	H	9	34.94	0.86	0.155

Table F.12: See caption of F.11.

TRIMER (β-ala)$_3$

HBs index	Donor NBO (i)					Acceptor NBO (j)					$\Delta E_{ij}^{(2)}$	$\epsilon_j^{(NL)} - \epsilon_i^{(L)}$	F_{ij}
1	BD	O	3	H	4	RY*	N	37	-	-	0.08	1.61	0.010
	BD	O	3	H	4	RY*	N	37	-	-	0.07	1.86	0.011
	BD	O	3	H	4	RY*	N	37	-	-	0.06	2.35	0.011
	BD	O	3	H	4	BD*	C	34	N	37	0.08	1.18	0.009
	BD	O	3	H	4	BD*	N	37	H	39	0.07	1.30	0.009
	BD*	O	3	H	4	RY*	N	37	-	-	1.07	0.02	0.014
	BD*	O	3	H	4	RY*	N	37	-	-	0.17	0.35	0.025
	BD	C	34	N	37	BD*	O	3	H	4	1.66	1.19	0.041
	BD	N	37	H	38	BD*	O	3	H	4	1.97	1.10	0.043
	BD	N	37	H	39	BD*	O	3	H	4	1.88	1.10	0.042
	CR	N	37	-	-	BD*	O	3	H	4	1.55	14.97	0.141
	LP	N	37	-	-	BD*	O	3	H	4	80.1	0.79	0.226
2	BD	C	15	O	18	BD*	N	37	H	39	0.31	0.96	0.015
	LP	O	18	-	-	BD*	N	37	H	39	2.53	1.31	0.051
	LP	O	18	-	-	BD*	N	37	H	39	1.48	0.85	0.033
	BD*	C	15	O	18	BD*	N	37	H	39	0.11	0.45	0.018
	BD	N	37	H	39	BD*	C	15	O	18	0.07	0.77	0.007
3	BD	C	21	N	24	BD*	O	29	H	30	0.15	1.26	0.012
	BD	N	24	H	25	BD*	O	29	H	30	0.14	1.16	0.012
	BD	N	24	H	26	BD*	O	29	H	30	1.45	1.15	0.037
	CR	N	24	-	-	BD*	O	29	H	30	0.23	15.02	0.054
	LP	N	24	-	-	BD*	O	29	H	30	26.84	0.86	0.136
	BD	O	29	H	30	RY*	N	24	-	-	0.07	1.78	0.010
	BD	O	29	H	30	BD*	C	21	N	24	0.07	1.23	0.008
	BD	O	29	H	30	BD*	N	24	H	25	0.18	1.33	0.014
	BD	O	29	H	30	BD*	N	24	H	26	0.1	1.33	0.010
4	BD	C	2	O	5	BD*	O	16	H	17	0.27	1.65	0.019
	BD	C	2	O	5	BD*	O	16	H	17	1.50	0.95	0.034
	CR	O	5	-	-	BD*	O	16	H	17	0.19	19.85	0.056
	LP	O	5	-	-	BD*	O	16	H	17	15.13	1.28	0.124
	LP	O	5	-	-	BD*	O	16	H	17	3.67	0.85	0.051
	BD*	C	2	O	5	BD*	O	16	H	17	0.66	0.43	0.038
	BD	O	16	H	17	RY*	O	5	-	-	0.12	1.63	0.013
	BD	O	16	H	17	RY*	O	5	-	-	0.36	2.54	0.027
	BD	O	16	H	17	BD*	C	2	O	5	0.34	1.51	0.020
	BD	O	16	H	17	BD*	C	2	O	5	0.25	0.9	0.014

Table F.13: See caption of F.11.

TETRAMER (β-ala)$_4$

HBs index	Donor NBO (i)					Acceptor NBO (j)					$\Delta E_{ij}^{(2)}$	$\epsilon_j^{(NL)} - \epsilon_i^{(L)}$	F_{ij}
1	BD	O	3	H	4	RY*	N	37	-	-	0.10	2.56	0.014
	BD	O	3	H	4	BD*	C	34	N	37	0.09	1.19	0.009
	BD	O	3	H	4	BD*	N	37	H	38	0.08	1.27	0.009
	BD*	O	3	H	4	RY*	N	37	-	-	0.09	0.40	0.020
	BD	C	34	N	37	BD*	O	3	H	4	1.22	1.20	0.035
	BD	N	37	H	38	BD*	O	3	H	4	1.53	1.10	0.038
	BD	N	37	H	39	BD*	O	3	H	4	2.06	1.10	0.044
	CR	N	37	-	-	BD*	O	3	H	4	1.33	14.97	0.131
	LP	N	37	-	-	BD*	O	3	H	4	71.11	0.80	0.214
2	BD	C	2	O	5	BD*	O	16	H	17	0.44	1.64	0.024
	BD	C	2	O	5	BD*	O	16	H	17	2.28	0.95	0.042
	CR	O	5	-	-	BD*	O	16	H	17	0.26	19.84	0.065
	LP	O	5	-	-	BD*	O	16	H	17	13.83	1.26	0.118
	LP	O	5	-	-	BD*	O	16	H	17	9.78	0.85	0.083
	BD*	C	2	O	5	BD*	O	16	H	17	1.06	0.42	0.047
	BD	O	16	H	17	RY*	O	5	-	-	0.05	2.58	0.011
	BD	O	16	H	17	BD*	C	2	O	5	0.19	1.50	0.015
	BD	O	16	H	17	BD*	C	2	O	5	0.44	0.91	0.019
3	BD	N	11	H	13	RY*	O	18	-	-	0.10	1.98	0.013
	BD	N	11	H	13	BD*	C	15	O	18	0.13	0.79	0.010
	BD	C	15	O	18	BD*	N	11	H	13	0.47	1.00	0.019
	LP	O	18	-	-	BD*	N	11	H	13	2.46	1.31	0.051
	LP	O	18	-	-	BD*	N	11	H	13	2.60	0.86	0.043
	BD*	C	15	O	18	BD*	N	11	H	13	0.20	0.44	0.024
4	BD	N	24	H	26	RY*	O	44	-	-	0.10	1.35	0.010
	BD	N	24	H	26	RY*	O	44	-	-	0.08	1.64	0.010
	BD	N	24	H	26	BD*	C	41	O	44	0.10	0.74	0.008
	BD	C	41	O	44	BD*	N	24	H	26	0.96	0.96	0.027
	LP	O	44	-	-	BD*	N	24	H	26	0.39	1.31	0.020
	LP	O	44	-	-	BD*	N	24	H	26	0.32	0.85	0.015
	BD*	C	41	O	44	BD*	N	24	H	26	0.11	0.45	0.018
5	BD	N	24	H	25	RY*	O	31	-	-	0.06	1.74	0.010
	BD	N	24	H	25	RY*	O	31	-	-	0.15	2.10	0.016
	BD	N	24	H	25	BD*	C	28	O	31	0.05	1.37	0.007
	BD	N	24	H	25	BD*	C	28	O	31	0.19	0.75	0.011
	BD	C	28	O	31	BD*	N	24	H	25	0.38	0.98	0.017
	LP	O	31	-	-	BD*	N	24	H	25	2.15	1.32	0.048
	LP	O	31	-	-	BD*	N	24	H	25	0.89	0.87	0.025
	BD*	C	28	O	31	BD*	N	24	H	25	0.16	0.46	0.021
6	BD	N	24	H	25	BD*	N	50	H	52	0.43	1.19	0.020
	BD	N	24	H	26	BD*	N	50	H	52	0.35	1.19	0.018
	LP	N	24	-	-	BD*	N	50	H	52	11.15	0.89	0.090
	BD	N	50	H	-	RY*	N	24	-	-	0.07	1.53	0.009
	BD	N	50	H	52	RY*	N	24	-	-	0.18	1.95	0.017
	BD	N	50	H	52	BD*	C	21	N	24	0.29	1.15	0.016

Table F.13: See caption of F.11.

TETRAMER (β-ala)$_4$

HBs index	Donor NBO (i)					Acceptor NBO (j)					$\Delta E_{ij}^{(2)}$	$\epsilon_j^{(NL)} - \epsilon_i^{(L)}$	F_{ij}
7	BD	O	29	H	30	RY*	N	50	-	-	0.06	1.71	0.009
	BD	O	29	H	30	RY*	N	50	-	-	0.08	2.63	0.013
	BD	O	29	H	30	BD*	C	47	N	50	0.06	1.21	0.008
	BD	O	29	H	30	BD*	N	50	H	51	0.08	1.29	0.009
	BD*	O	29	H	30	RY*	N	50	-	-	0.07	0.46	0.022
	BD*	O	29	H	30	RY*	N	50	-	-	0.07	0.33	0.018
	BD*	O	29	H	30	RY*	N	50	-	-	0.09	0.26	0.018
	BD*	O	29	H	30	BD*	N	50	H	51	0.09	0.05	0.008
	BD	C	47	N	50	BD*	O	29	H	30	0.75	1.23	0.028
	BD	N	50	H	51	BD*	O	29	H	30	0.85	1.12	0.028
	BD	N	50	H	52	BD*	O	29	H	30	1.92	1.13	0.043
	CR	N	50	-	-	BD*	O	29	H	30	0.83	14.99	0.102
	LP	N	50	-	-	BD*	O	29	H	30	53.91	0.82	0.189
8	BD	C	8	N	11	BD*	O	42	H	43	0.90	1.21	0.030
	BD	N	11	H	12	BD*	O	42	H	43	1.26	1.11	0.034
	BD	N	11	H	13	BD*	O	42	H	43	1.96	1.11	0.043
	CR	N	11	-	-	BD*	O	42	H	43	1.07	14.98	0.117
	LP	N	11	-	-	BD*	O	42	H	43	64.44	0.81	0.205
	BD	O	42	H	43	RY*	N	11	-	-	0.11	1.77	0.013
	BD	O	42	H	43	RY*	N	11	-	-	0.09	2.57	0.013
	BD	O	42	H	43	BD*	C	8	N	11	0.12	1.21	0.011
	BD*	O	42	H	43	RY*	N	11	-	-	0.06	0.17	0.011
	BD*	O	42	H	43	RY*	N	11	-	-	0.47	0.13	0.027
	BD*	O	42	H	43	BD*	N	11	H	12	0.06	0.07	0.007

Table F.14: See caption of F.11.

PENTAMER (β-ala)$_5$

HBs index	Donor NBO (i)					Acceptor NBO (j)					$\Delta E_{ij}^{(2)}$	$\epsilon_j^{(NL)} - \epsilon_i^{(L)}$	F_{ij}
1	BD	O	16	H	17	RY*	N	50	-	-	0.05	1.72	0.008
	BD	O	16	H	17	RY*	N	50	-	-	0.08	2.31	0.012
	BD	O	16	H	17	BD*	N	50	H	51	0.09	1.30	0.010
	BD*	O	16	H	17	RY*	N	50	-	-	0.14	0.38	0.029
	BD*	O	16	H	17	RY*	N	50	-	-	0.07	0.24	0.016
	BD*	O	16	H	17	BD*	N	50	H	51	0.08	0.05	0.008
	BD	C	47	N	50	BD*	O	16	H	17	0.54	1.24	0.024
	BD	N	50	H	51	BD*	O	16	H	17	0.59	1.13	0.024
	BD	N	50	H	52	BD*	O	16	H	17	1.38	1.14	0.036
	CR	N	50	-	-	BD*	O	16	H	17	0.59	15.00	0.086
	LP	N	50	-	-	BD*	O	16	H	17	44.90	0.84	0.174
2	BD	N	50	H	52	RY*	N	63	-	-	0.11	1.72	0.012
	BD	N	50	H	52	RY*	N	63	-	-	0.10	2.44	0.014
	BD	N	50	H	52	BD*	C	60	N	63	0.11	1.13	0.010
	BD	N	63	H	64	BD*	N	50	H	52	0.61	1.18	0.024
	LP	N	63	-	-	BD*	N	50	H	52	10.35	0.89	0.086

Table F.14: See caption of F.11.

PENTAMER (β-ala)$_5$

HBs index	Donor NBO (i)					Acceptor NBO (j)					$\Delta E_{ij}^{(2)}$	$\epsilon_j^{(NL)} - \epsilon_i^{(L)}$	F_{ij}
3	BD	C	15	O	18	BD*	N	63	H	64	1.55	0.99	0.035
	LP	O	18	-	-	BD*	N	63	H	64	2.65	1.32	0.053
	LP	O	18	-	-	BD*	N	63	H	64	0.64	0.87	0.022
	BD*	C	15	O	18	BD*	N	63	H	64	0.45	0.46	0.037
	BD	N	63	H	64	BD*	C	15	O	18	0.30	0.76	0.014
4	BD	O	3	H	4	RY*	O	44	-	-	0.14	1.95	0.015
	BD	O	3	H	4	RY*	O	44	-	-	0.28	2.52	0.024
	BD	O	3	H	4	BD*	C	41	O	44	0.45	1.51	0.023
	BD	O	3	H	4	BD*	C	41	O	44	0.26	0.92	0.015
	BD	C	41	O	44	BD*	O	3	H	4	0.42	1.64	0.024
	BD	C	41	O	44	BD*	O	3	H	4	0.60	0.96	0.022
	CR	O	44	-	-	BD*	O	3	H	4	0.27	19.85	0.066
	LP	O	44	-	-	BD*	O	3	H	4	20.42	1.27	0.144
	LP	O	44	-	-	BD*	O	3	H	4	2.09	0.84	0.038
	BD*	C	41	O	44	BD*	O	3	H	4	0.70	0.42	0.039
5	BD	C	8	N	11	BD*	O	55	H	56	0.41	1.25	0.021
	BD	N	11	H	12	BD*	O	55	H	56	0.41	1.16	0.020
	BD	N	11	H	13	BD*	O	55	H	56	1.84	1.15	0.042
	CR	N	11	-	-	BD*	O	55	H	56	0.53	15.02	0.081
	LP	N	11	-	-	BD*	O	55	H	56	41.64	0.86	0.169
	BD	O	55	H	56	RY*	N	11	-	-	0.05	1.66	0.008
	BD	O	55	H	56	RY*	N	11	-	-	0.06	1.50	0.009
	BD	O	55	H	56	BD*	C	8	N	11	0.07	1.22	0.008
	BD	O	55	H	56	BD*	N	11	H	12	0.13	1.32	0.012
6	BD	C	2	O	5	BD*	N	37	H	38	0.91	0.98	0.027
	LP	O	5	-	-	BD*	N	37	H	38	1.63	1.30	0.041
	LP	O	5	-	-	BD*	N	37	H	38	2.29	0.85	0.041
	BD*	C	2	O	5	BD*	N	37	H	38	0.21	0.43	0.024
	BD	N	37	H	38	BD*	C	2	O	5	0.11	0.77	0.009
7	BD	C	34	N	37	BD*	O	42	H	43	1.84	1.15	0.043
	BD	N	37	H	38	BD*	O	42	H	43	4.48	1.06	0.064
	BD	N	37	H	39	BD*	O	42	H	43	2.02	1.06	0.043
	CR	N	37	-	-	BD*	O	42	H	43	2.16	14.93	0.168
	LP	N	37	-	-	BD*	O	42	H	43	93.72	0.76	0.239
	BD	O	42	H	43	BD*	C	34	N	37	0.16	1.19	0.012
	BD	O	42	H	43	BD*	N	37	H	39	0.10	1.28	0.010
	BD*	O	42	H	43	RY*	N	37	-	-	0.06	1.66	0.031
	BD*	O	42	H	43	BD*	C	34	N	37	1.73	0.04	0.024
	BD*	O	42	H	43	BD*	N	37	H	39	0.21	0.13	0.015
8	BD	C	21	H	22	BD*	O	29	H	30	0.08	1.02	0.008
	BD	C	21	N	24	BD*	O	29	H	30	0.99	1.25	0.032
	BD	C	21	N	24	BD*	O	29	H	30	0.99	1.25	0.032
9	BD	N	24	H	25	BD*	O	29	H	30	1.28	1.16	0.035
	BD	N	24	H	26	BD*	O	29	H	30	0.27	1.16	0.016
	CR	N	24	-	-	BD*	O	29	H	30	0.58	15.03	0.085
	LP	N	24	-	-	BD*	O	29	H	30	42.86	0.86	0.172
	BD	O	29	H	30	RY*	N	24	-	-	0.11	1.83	0.013
	BD	O	29	H	30	BD*	N	24	H	26	0.22	1.31	0.015

Appendices
List of publications included as part of the thesis

1 S. Maclot, D. G. Piekarski, A. Domaracka, A. Méry, V. Vizcaino, L. Adoui, F. Martín, M. Alcamí, B. A. Huber, P. Rousseau and S. Díaz-Tendero,
"Dynamics of glycine dications in the gas phase: ultrafast intramolecular hydrogen migration versus Coulomb repulsion",
The Journal of Physical Chemistry Letters 5, 3903 (2013).

2 S. Maclot, D. G. Piekarski, R. Delaunay, A. Domaracka, A. Méry, V. Vizcaino, J-Y. Chesnel, F. Martín, M. Alcamí, B. A. Huber, L. Adoui, P. Rousseau and S. Díaz-Tendero,
"Stability of the glycine cation in the gas phase after interaction with multiply charged ions", **European Physical Journal D** 68, 149 (2014).

3 D. G. Piekarski, R. Delaunay, S. Maclot, L. Adoui, F. Martín, M. Alcamí, B. A. Huber, P. Rousseau, A. Domaracka and S. Díaz-Tendero,
"Unusual hydroxyl migration in the fragmentation of β-alanine dication in the gas phase",
Physical Chemistry Chemical Physics 17, 16767, (2015).

4 J. Kocisek, D. G. Piekarski, R. Delaunay, B. A. Huber, L. Adoui, F. Martín, M. Alcamí, P. Rousseau, A. Domaracka, J. Kopyra, and S. Díaz-Tendero,
"N-acetylglycine cation tautomerization enabled by the peptide bond",
The Journal of Physical Chemistry A 119, 9581 (2015).

5 S. Maclot, R. Delaunay, D. G. Piekarski, A. Domaracka, B. A. Huber, L. Adoui, F. Martín, M. Alcamí, L. Avaldi, P. Bolognesi, S. Díaz-Tendero and P. Rousseau,
"Determination of Energy-Transfer Distributions in Ionizing Ion-Molecule Collisions",
Physical Review Letters 117, 073201 (2016).

6 D. G. Piekarski and S. Díaz-Tendero,
"Structure and stability of clusters of β-alanine in the gas phase: importance of the nature of intermolecular interactions",
Physical Chemistry Chemical Physics 19, 5465 (2017).

7 D. G. Piekarski, R. Delaunay, A. Mika, S. Maclot, L. Adoui, F. Martín, M. Alcamí, B. A. Huber, P. Rousseau, S. Díaz-Tendero and A. Domaracka,
"Production of doubly-charged highly reactive species from the long-chain amino acid GABA initiated by Ar^{9+} ionization."
Physical Chemistry Chemical Physics,
accepted (2017), DOI:10.1039/C7CP00903H.

Appendix H

Appendices
List of publications not included as part of the thesis

1 P. López-Tarifa, D. G. Piekarski, E. Rossich, M.-A. Hervé Du Penhoat, R. Vuilleumier, M.-P. Gaigeot, I. Tavernelli, M.-F. Politis, Y. Wang, S. Díaz-Tendero, F. Martín and M. Alcamí,
"Ultrafast nonadiabatic fragmentation dynamics of biomolecules",
Journal Physics: Conference Series 488, 012037 (2014).

2 E. Kukk, D. T. Ha, Y. Wang, D. G. Piekarski, S. Díaz-Tendero, K. Kooser, E. Itälä, H. Levola, M. Alcamí, E. Rachlew and F. Martín, *"Internal energy dependence in x-ray-induced molecular fragmentation: An experimental and theoretical study of thiophene"*,
Physical Review A, 91, 043417 (2015).

3 *"Stability, structure and peptide bond formations in the ionised clusters of β-alanine molecules"*, to be submitted.

4 *"Surface hopping nonadibatic molecular dynamics with complete active space study of betaine"*, to be submitted.

5 *"Surface hopping with linear response time-dependent density functional theory for S-α-alanine"*, to be submitted.

6 *"Born-Oppenheimer molecular dynamics study with potential energy surface study of the nano-solvated Br-uracil"*, to be submitted.